



高新科技译丛

Bayesian Estimation and Tracking
A practical Guide

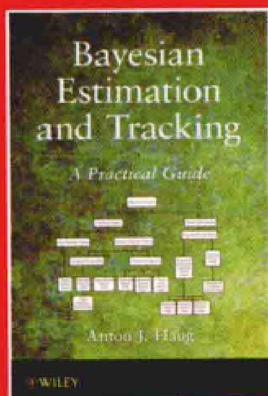
贝叶斯估计与跟踪实用指南

【美】 Anton J.Haug 著　　王欣 于晓 译

WILEY



国防工业出版社
National Defense Industry Press



Bayesian Estimation and Tracking
A practical Guide

内容简介

非高斯系统估计的文献一般都存在缺少实用方法的缺点，而高斯方法又缺乏严密、有组织的推导。《贝叶斯估计与跟踪实用指南》一书设法解决这个领域中的不足，为读者提供全面的高斯噪声与非高斯噪声驱动的线性与非线性动态系统估计方法的介绍。

本书特点在于提供了贝叶斯估计与跟踪的统一方法，重点强调了在贝叶斯框架下所有跟踪算法的推导过程，还给出了用于计算加权密度积分的高效的数值方法，包括加权高斯积分的线性与非线性卡尔曼滤波器，以及非高斯情况下的粒子滤波器。作者首先给出每种估计方法基本原理的详细推导过程，然后分步骤对每一种方法进行详细说明，使得跟踪滤波器的编码更简单易懂。

本书利用了案例分析来展示所研究内容的应用。此外，该书还为每个算法提供了框图，让读者可以自己开发MATLAB®估计方法软件工具箱。

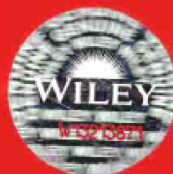
《贝叶斯估计与跟踪实用指南》是一本优秀的估计与跟踪方法课程的研究生教材。本书还适用于估计与跟踪领域的科学研究人员、数学研究人员和工程技术人员，对他们具有重要的实际应用价值。对广大科技工作者来说，也是一本非常有价值的参考书。

作者简介

安东.J.霍格，博士，美国约翰.霍普金斯大学应用物理实验室科研人员，他为防空和导弹防御系（AMDD）提出了先进的目标跟踪方法。霍格博士的研究领域包括目标跟踪，信号和阵列处理与处理设计，主动、被动雷达和声纳设计，数字通信和编码理论以及时间频率分析。

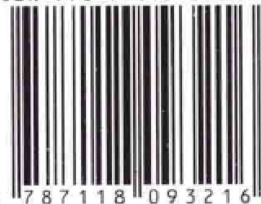
责任编辑：牛旭东　xdniu@ndip.cn
责任校对：苏向颖
封面设计：蒋秀芹

WILEY

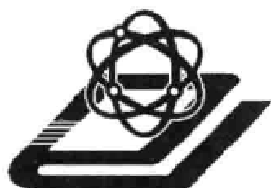




▶ 上架建议：计算机视觉 ◀
http://www.ndip.cn



定价：79.00 元



贝叶斯估计与跟踪实用指南

Bayesian Estimation and Tracking: A Practical Guide

[美] Anton J. Haug 著
王欣 于晓 译

国防工業出版社
·北京·

著作权合同登记　图字:军-2013-200号

图书在版编目(CIP)数据

贝叶斯估计与跟踪实用指南/(美)豪格(Hang,A. J.)著;王欣,于晓译.—北京:国防工业出版社,2014.5
(高新科技译丛)
书名原文:Bayesian estimation and tracking:a practical guide
ISBN 978-7-118-09321-6

Ⅰ.①贝… Ⅱ.①豪… ②王… ③于… Ⅲ.①贝叶斯估计-指南②贝叶斯估计-应用-目标跟踪-指南Ⅳ.①0211.67-62②TN953-62

中国版本图书馆CIP数据核字(2014)第068301号

Authorized translation from the English language edition, entitled Bayesian Estimation and Tracking, ISBN 978-0-470-62170-7, by Anton J. Haug, published by John Wiley & Sons.

※

国防工業出版社出版发行
(北京市海淀区紫竹院南路23号　邮政编码100048)
北京嘉恒彩色印刷有限责任公司
新华书店经售
*
开本 710×1000　1/16　**印张** 19¾　**字数** 370千字
2014年5月第1版第1次印刷　**印数** 1—2000册　**定价** 79.00元

(本书如有印装错误,我社负责调换)

国防书店:(010)88540777　发行邮购:(010)88540776
发行传真:(010)88540755　发行业务:(010)88540717

前　言

本书以基本原理为出发点，阐述了贝叶斯估计滤波器的完整的发展历程。书中研究了由高斯噪声和非高斯噪声驱动的线性、非线性动态系统。这里假设动态系统是连续的，因为只有在研究离散滤波器的时候，才会在离散时间对系统进行观测。本书的主要目标是给读者提供过去60年来所提出的大部分贝叶斯估计方法的一致性的全面描述，以表明每种以贝叶斯范式为基础的基本思想是如何与条件密度相关的。

学习本书内容需要的预备知识包括线性代数、贝叶斯概率理论以及有限差分和插值的数值方法。书中第2章介绍了所有这些主题以及掌握书中其它内容所需的基础知识。

书中涵盖的部分内容来自于马里兰大学帕克分校数学系的一门研究生课程。该门课程的主要目的是让学生自己开发目标跟踪算法的MATLAB工具箱。课后给学生们提出一些特定的跟踪问题，要求编写一个或多个MATLAB子程序来实现某些跟踪方法。一般来说，学生开发的子程序是独立的，并且可以应用于多种难度不大的跟踪问题（例如，纯方位跟踪）。本书不包含任何习题集，使用本书的读者可以根据需要来调整跟踪问题的作业。此外，本书包含4个相当复杂的实用案例，使用本书的老师可以选择其中之一作为编码作业的框架。第一个案例贯穿本书的第Ⅱ和第Ⅲ部分；其余的3个案例则被分成了单独的章节，构成了本书的第Ⅳ部分。

本书有两个鲜明特点：①运用大量的表格和图形，给出每种估计方法基本原理的详细推导过程；②每种方法都有非常详细的分步指导，使得跟踪滤波器的编码简单且易于理解。

递推贝叶斯估计可以求解转移函数或变换函数的加权条件密度积分。转移函数反映动态状态向量随时间变化的路径，变换函数将状态向量变换成观测向量。有多种求解积分的数值方法，每种都有不同的估计方法。本书第Ⅱ部分和第Ⅲ部分的每个章节都涵盖求解这些积分的一个或多个数值逼近方法，第Ⅱ部分是高斯加权积分的卡尔曼滤波器方法，第Ⅲ部分是未知密度的加权密度积分的粒子滤波器方法。

本书是作者在此领域多年研究的成果，希望能为贝叶斯估计和跟踪研究做出贡献。同时，也希望在研究改进这些方法并将其应用到新领域、解决新问题的基础上，本书能够开拓贝叶斯估计和跟踪研究的新方向。

Anton J. Haug

目　　录

第Ⅲ部分 蒙特卡洛方法

BAYESIAN

第 I 部分

预备知识

第 1 章　简介
第 2 章　数学基础
第 3 章　贝叶斯估计的基本概念
第 4 章　实用案例:初步探讨

第1章 简 介

自统计数学出现以来,动态系统估计与跟踪已成为许多数学家的研究热点。过去50年中已提出许多估计方法,用于对线性和高斯分布动态系统进行统计推理(估计)。对于非线性高斯系统的应用,有些方法以增加计算复杂度为代价,也得了很好的估计结果。然而,现实世界的动态系统中,无论是线性的还是非线性的,通常有异常值过多的现象,并呈现为非高斯分布。已经证明,在处理这种真实世界数据的时候,常规的高斯估计方法不足以处理这些问题,会导致估计滤波器无法收敛。

随着高速桌面计算的出现,在过去的十年中,数学研究的重点已经转移到非高斯分布动态系统之上。但是,大部分的非高斯系统估计相关研究成果都是理论上的,那种普通工程人员无需太多学习就可以方便使用的方法基本没有。已提出的用于非线性系统估计的高斯方法,都缺少严密、有组织的推导。在概念上,高斯方法和非高斯方法仍然缺乏连续性。

本书致力于将最新的动态系统统计估计方法的研究成果全面地呈现给读者,包括线性、非线性系统和高斯、非高斯的系统等。利用贝叶斯框架,给出一个概念上紧密相关的路线图。首先,从基本原理开始介绍,进而引出一些目前广泛使用的高斯估计方法;然后,将这些概念推广到非高斯估计领域,并采用蒙特卡洛方法进行估计。虽然贝叶斯方法可求得统计密度估计,但多数情况下使用点估计方法,计算加权密度积分并得统计密度估计。因此,本书主要介绍在高斯、非高斯密度下的加权密度积分数值计算方法。

本书还为每种估计算法提供了伪代码和框图,供读者开发跟踪软件工具箱。为帮助读者理解书中所给出的算法,本书还将针对几个非常具体的实际问题,介绍一些开发仿真所需要的知识。书中将给出案例分析,并给出每种跟踪估计的结果。因为很难专门地去对跟踪方法进行比较,所以本书给出了多种评价方法,用

于评价各种跟踪滤波器的相对性能。

1.1 贝叶斯推理

推理方法是对一组参数估计出当前值的过程,这组参数基于一组观察值或测量值。估计过程一般遵循以下两个模型。第一个模型假定待估计参数(通常是不可观测的)是非随机的并且在观察窗中保持恒定,但是观测值含有噪声和随机分量;第二个模型假定参数是随机变量,具有一定先验概率,且观测含噪。利用第一个模型进行参数估计的过程被称为非贝叶斯估计或 Fisher 估计[1],利用第二个模型进行参数估计被称为贝叶斯估计。

贝叶斯估计的概念非常简单,它基于一些初始先验信念。例如:"注意那艘船。它离海岸 914m(1000yd),向东北方向以 10n mile/h 的速度运动。"初始信念的陈述(包括最初猜测的船舶的位置和速度)是不确定或随机的,且基于某些先验的概率分布。基于初始信念,可以做出如下预测:由于船舶以恒定速度运动,它将在大约 10min 之后到达那个地方。这个声明包括船舶运动动力学概念模型以及一些不确定性。现在假设,一个人手上有一个小的便携式雷达,该雷达能够以一定的精度测量(观测)船舶视线范围内的距离和速度。给定适当的数学模型,可将笛卡儿坐标系下观测到的船舶位置和速度联系起来,当前的雷达测量值可以用来更新预测船舶状态(位置与速度)。

上面一段阐述了递归贝叶斯估计的实质:

(1)给出初始信念陈述;

(2)利用先验信念和动态模型进行预测;

(3)利用观测值和观测模型来得后验信念,更新预测值;

(4)将后验信念置成新的先验信念,返回步骤(2)。

递推估计的概念最初是在神甫托马斯·贝叶斯的一篇论文中首次提出的,而在贝叶斯去世数年之后(1763 年),由理查德·普莱斯投稿到皇家统计学会。在 Stephen Brooks[2] 的论文中,可以读到与贝叶斯统计推理有关的历史和思想的精彩回顾。Brooks 的论文中还有一些值得关注的例子,例如进行了贝叶斯方法与所谓"频率论者"的统计推理方法之间的对比。因为本书致力于研究贝叶斯方法,不再深入探讨频率论的方法,感兴趣的读者可参考 Brooks 的论文。

1.2 贝叶斯层次估计法

如上所述,本书将对现代跟踪滤波器中一个子集进行严密有组织的推导。图 1-1 给出了本书提出的跟踪滤波器层次结构。

图 1-1 左边是高斯跟踪滤波器,右边是蒙特卡洛非高斯滤波器。在后续章

节对每种跟踪滤波器进行详细阐述的过程中，将以此图作为向导，通过它来定位学习的进程。书中还会介绍其他一些感兴趣的概念，如讨论如何评价性能等。在多数情况下，书中将按照从上到下、由左至右的方式，对框图内容加以介绍。在图 1 – 1 中右下角内容介绍完毕之后，读者将会对所有的贝叶斯跟踪滤波器之间的关联有一个全面的认识。

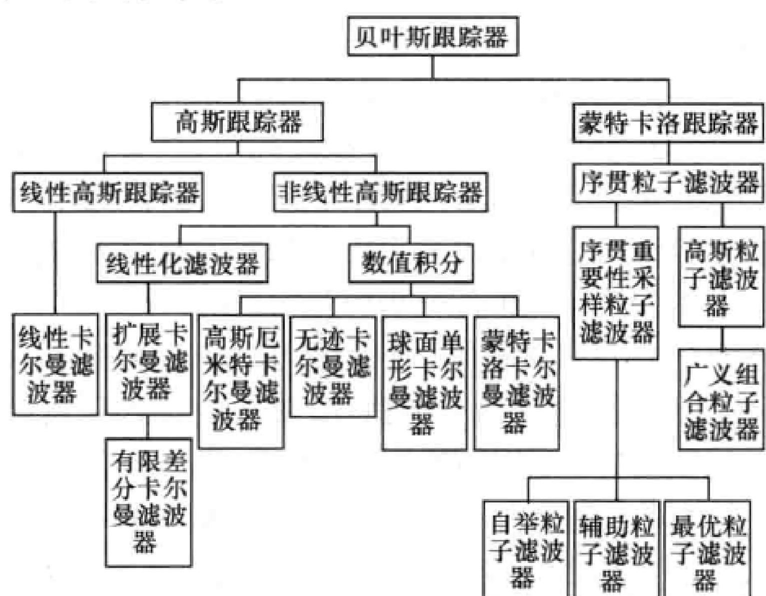


图 1 – 1　贝叶斯估计的跟踪滤波器的层次结构

1.3　本书研究范畴

1.3.1　目标

本书主要目的是让读者能够清晰地了解各种贝叶斯估计方法及其相互关系。首先从贝叶斯理论的基本原理开始介绍，并在先前的滤波器基础上，通过细微的改进推导出各种跟踪滤波器。这种推导方式可以让读者更深入地理解贝叶斯估计与跟踪算法的层次关系。在介绍跟踪滤波器的过程中，为每种滤波器提供了伪代码和程序流程图，以方便读者查询及参考。

在 Jazwinski[3] 最早出版于 1970 年的关于滤波理论的书中，有如下论述："本书的意义是双重的。首先，尽管线性估计理论广为人知，但多数成果都分散在期刊文献当中，没有被搜集在一起；其次，现有文献在连续非线性理论上的研究非常深奥、争议很大。因此，如果对测度理论和随机微分方程没有深入了解，普通工程师很难将这些理论应用到实际当中。"类似的论述也适用于估计理论中非高斯蒙特卡洛算法的研究现状。大多数已发表的成果都非常深奥，对测度理论不了解的普通工程师无法将其应用到实践中。Doucet 等[4] 的特邀论文集就是一个典型的例子，在这部优秀的特邀论文集中，许多章节都非常深奥。

本书将沿用 Jazwinski 的方法，删减过多的基于测度理论的阐述，普通工程师很难理解这些深奥的数学知识，希望本书没有采用同样深奥的数学理论来替代它们。

1.3.2 章节概述

本书不是一本入门级的教科书，其编写目的是作为一门研究生课程，或者给希望深入学习跟踪估计算法的读者提供参考。本书的读者需要对概率论相关的书籍（如 Papoulis[5]）达到研究生水平的理解。同时，读者还应该熟练掌握矩阵线性代数和数值方法，包括有限差分。为减少读者学习本书的困难，第 2 章介绍了贯穿全书的数学工具。对这部分内容可以有选择性地阅读，有经验的读者可以跳过这章，对于学习掌握第 Ⅱ、Ⅲ 部分内容没有影响。

第 Ⅰ 部分分为4 章。第2 章概括介绍了学习估计算法所需要的数学理论。第 3 章介绍了动态系统贝叶斯递归估计的概念，该系统可以用离散隐式马尔可夫过程进行建模。在观测模型中，观测值为离散的，与系统状态相关。本书一般不考虑连续估计算法。第4 章是案例分析 —— 通过定向频率分析与记录（DIFAR）声纳浮标得分布场，进行船舶跟踪。案例分析中包括给出系统动力学的全仿真和生成含噪观测值等。

第 Ⅱ 部分涵盖了高斯噪声情况下估计方法的提出与应用。第 5 章用高斯概率密度对第3 章提出的一般贝叶斯估计方法进行了修改。在第6 章 ~ 第12 章，对高斯卡尔曼滤波器算法进行了推导和代码实现，包括线性卡尔曼滤波器（LKF）、扩展卡尔曼滤波器（EKF）、有限差分卡尔曼滤波器（FDKF）、无迹卡尔曼滤波器（UKF）、球面单形卡尔曼滤波器（SSKF）、高斯厄米特卡尔曼滤波器（GHKF）和蒙特卡洛卡尔曼滤波器（MCKF）。除 MCKF 以外，后 5 种跟踪滤波器中的 4 种都可以归类为一般 Sigma 点卡尔曼滤波器，即：利用确定性矢量点积分来计算高斯加权积分，并进一步用来估计状态矢量的均值和协方差矩阵。MCKF 用采样分布代替连续高斯分布，这样估计积分运算就简化为求和运算，保持了非线性函数的特性。在第 13 章中可以看到，后 5 种卡尔曼滤波方法可以概括为一种估计方法，它们的区别仅仅是用到的矢量点的数量、位置和权重的变化。

在很多估计算法文献中，常常忽略了一个关键问题 —— 估计算法性能的量化测度问题。本书第 14 章的主题是研究算法测度问题，给出在真实系统环境下，基于估计协方差矩阵计算和测绘误差椭圆的方法，以及蒙特卡洛仿真环境下，计算均方根（RMS）误差、Cramer - Rao 下界（CRLB）的方法。第 14 章最后一节主要研究各种估计方法在 DIFAR 浮标跟踪中的应用，以及降低输入信号信噪比时的性能对比结果。

第 Ⅲ 部分详细讨论基于非高斯概率分布的估计算法。在第 12 章中，采用一组离散蒙特卡洛采样来实现高斯密度估计，将求均值、求协方差中估计积分值的

运算简化为求加权和运算，分别称为采样均值和采样协方差。高斯密度的采样权值通常为 $1/N$，这里 N 代表采样点数目。非高斯密度采样有如下两个问题：① 从密度函数中直接得蒙特卡洛采样非常困难；② 当密度函数不存在一阶矩和二阶矩的时候，采样会出现问题，柯西密度是一个典型的例子。为解决上述采样问题，第 15 章引入了蒙特卡洛方法并提出了重要性采样的概念，在此基础上推导出粒子滤波器估计方法（这里的粒子就是蒙特卡洛采样点）。在粒子滤波器使用过程中，会出现一些问题，为解决这些问题，研究者提出了很多改进方法。对于重要性采样，每个采样点的权值都是通过计算该采样点的非高斯密度与重要性密度的比值得的。在一定的假设条件下，可以递归地计算权重，这样就产生了序贯重要性采样（SIS）粒子滤波器（详见第16章）。第17章提出在滤波器每一步迭代过程中重新计算权值，由此产生了高斯组合粒子滤波器。在第 17 章的末尾，还给出 DIFAR 应用中，各种粒子滤波跟踪估计方法性能的比较结果。

几本最近出版的书籍为书中研究的主题提供了一些有益的参考。其中，在学习高斯卡尔曼滤波器时，Bar Shalom 等[6] 和 Candy[7] 的文献是非常好的指南书籍。而对于学习第 Ⅲ 部分的非高斯滤波方法来说，Doucet 等和 Ristic 等[8] 的著作则是非常优秀的参考书。

1.4 用 MATLAB® 进行建模和仿真

学习过程非常重要的是能够为读者提供一些具体实例，来实现复杂问题的估计。本书通过使用MATLAB® 进行仿真来保证这一点。本书共提供4个案例分析，从第一个到最后一个，复杂度逐渐增加。每一个案例分析中，首先介绍如何对系统进行动态建模仿真和观测建模仿真；然后说明如何用蒙特卡洛方法生成随机观测数据；之后运用各种跟踪滤波器对仿真的观测数据进行处理，产生多组跟踪数据；最后用跟踪数据对各种跟踪滤波器加以比较。

第一个案例研究的是船舶跟踪问题。当运动船舶穿过 DIFAR 浮标分布场时，对它进行跟踪。DIFAR 浮标利用船舶发射的宽带含噪信号作为输入，将得的船舶方位含噪观测值作为输出。第4 章内容表明，DIFAR 浮标输出的方位估计概率密度依赖于输入信号的信噪比，即：信噪比很高时，概率密度为高斯分布，随着信噪比不断减小，概率密度会逐渐变为均匀分布。研究这一案例的目的在于，当观测噪声从高斯分布变为非高斯分布时，检测每种跟踪估计算法的滤波跟踪性能。

本书利用DIFAR实例，依次说明各种跟踪估计滤波器。第4章将介绍如何实现 DIFAR 浮标的仿真，以产生模拟的、信噪比相关的观测值集合。在后续章节中，采用上述方位观测值集合，运用各种跟踪算法来产生蒙特卡洛跟踪估计集合，清晰展现高斯观测噪声变为非高斯观测噪声对各种跟踪方法的影响。

第 Ⅳ 部分还给出另外三个案例分析，用来说明第 Ⅱ 和第 Ⅲ 部分介绍的跟踪滤波器的使用方法。

第 18 章主要解决如何实现三维空间机动目标跟踪这一关键问题。这里介绍了一种新方法，使用恒定球面速度模型来取代经典的恒定笛卡儿速度模型，并阐释在使用多种高斯跟踪滤波器跟踪机动目标时，球面模型在改善跟踪性能上的表现。

第 19 章给出了第 3 个案例分析，其中研究了相当复杂的下落炸弹动态跟踪问题，具体是使用多跟踪点视频帧来实现跟踪，跟踪点包括投下炸弹的飞机和炸弹本身。这是一种被称为摄影测量法的复杂过程的案例，对象的几何属性和运动学属性都可以从连续拍摄的图像帧中推算得，这个案例中包含非常复杂的非线性多维观测过程和非线性多维动力学模型。而且，这里动力学模型和观测模型都是高维的，是跟踪滤波器中最为复杂的问题。这里通过维度的影响来证明，使用所有的跟踪滤波器在计算上是不可行的。

第 20 章给出最后一个案例分析，改进了摄影测量法在估计中的应用，使用多个传感器，如多个摄像头、平移加速度计和角速率陀螺仪等，将多传感器采集到的数据进行融合。这些数据在单独使用时，每种数据源都有它的优势，但也都有缺点。当使用多个不同的传感器联合估计时，得的估计结果通常更为准确、可靠。通过对估计结果的分析表明，传感器融合辅助估计器可以从飞行测试中恢复有意义的结果，否则测试就会失败。

参考文献

[1] Fisher R. *Statistical Methods and Scientific Inference.* Revised Edition. Macmillan Pub. Co,1973.

[2] Brooks SP. Bayesian computation: a statistical revolution. *Phil. Trans. R. Soc. Lond. A*, 2003,361: 2681 - 2697.

[3] Jazwinski AH. *Stochastic Processes and Filtering Theory.* 2th ed. Dover Publications,2007.

[4] Doucet A,de Freitas JFG,Gordon NJ. *Sequential Monte Carlo Methods in Practice.* New York,NY: Springer - Verlag,2001.

[5] Papoulis A. *Probability,Random Variables,and Stochastic Processes.* 4th ed. McGraw - Hill,2002.

[6] Bar Shalom Y,Li XR,*Kirubarajan T. Estimation with Application to Tracking and Navigation: Theory, Algorithms and Software.* Wiley, 2001.

[7] Candy JV. *Bayesian Signal Processing: Classical,Modern,and Particle Filtering Methods.* Hoboken,NJ: Wiley,2009.

[8] Ristic B,Arulampalam S,Gordon N. *Beyond the Kalman Filter: Particle Filters for Tracking Applications.* Boston,MA: Artech House,2004.

第2章 数学基础

本章将介绍一些数学概念，它们是理解第 Ⅱ 至 Ⅳ 部分内容必不可少的知识。这里仅将对相关数学知识进行简要介绍，而不会涉及大量的、严格的证明。对于这里所涵盖的大多数主题，将假定读者很熟悉之前提到的数学概念。虽然不建议，但读者可以在初次阅读的时候跳过这个章节，在阅读第 Ⅰ 部分其他内容和第 Ⅱ 至 Ⅳ 部分时，可将本章作为索引使用。补充的其它数学概念将会根据需要加以介绍。

2.1 矩阵线性代数简介

本节给出本书中所用到的符号约定和矩阵运算，并利用 MATLAB® 环境，提供一些矩阵线性代数概念的简要概况。在线的参考文献[1] 对于本节介绍的矩阵运算和公式是非常好的参考资料。

2.1.1 矢量与矩阵的约定与记法

通常用小写字母来表示标量，例如 a；用小写的黑斜体字母来表示一个矢量，例如 $\boldsymbol{a}$；用大写的黑斜体字母来表示矩阵，例如 $\boldsymbol{A}$。

一个 n 维矢量是 $n \times 1$ 的列矢量，即

$$\boldsymbol{a} = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} \tag{2-1}$$

按照惯例，所有的矢量都是列矢量。可以将列矢量写成行矢量的转置（在报告或期刊中，有利于节省空间），即

$$\boldsymbol{a} = [a_1, \cdots, a_n]^{\mathrm{T}} \tag{2-2}$$

用逗号隔开矢量中的元素。$[\cdot]^{\mathrm{T}}$ 表示一个矢量或者矩阵的转置，后面将给出其定义。

一个 $n\times m$ 矩阵可表示为如下的二维数组，即

$$\boldsymbol{A}=\begin{bmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nm} \end{bmatrix} \tag{2-3}$$

式中：第一个下标是元素所在行的索引；第二个下标是元素所在列的索引。矩阵 $\boldsymbol{A}$ 的第 ij 个元素是 a_{ij}。

用符号 $\boldsymbol{A}^{\mathrm{T}}$ 表示矩阵 $\boldsymbol{A}$ 的转置，定义为

$$\boldsymbol{A}^{\mathrm{T}}=\begin{bmatrix} a_{11} & \cdots & a_{n1} \\ \vdots & \ddots & \vdots \\ a_{1m} & \cdots & a_{nm} \end{bmatrix} \tag{2-4}$$

$n\times n$ 方阵 $\boldsymbol{A}$ 的迹被定义为矩阵 $\boldsymbol{A}$ 的主对角线（从左上到右下的对角线）上元素的和，即

$$\mathrm{tr}\boldsymbol{A}=\sum_{i=1}^{n} a_{ii} \tag{2-5}$$

同时，矩阵的迹也是其特征值的总和，因此其具有基变换下的不变性。这一特性可以用来定义一般的线性算子的迹。

本书偶尔也会使用符号 $\boldsymbol{A}\in\mathbf{R}^{n\times m}$ 或者 $\boldsymbol{a}\in\mathbf{R}^{n}$ 来分别指定矩阵或者矢量的维度。

2.1.2 和与乘积

矩阵的加法和数乘定义为

$$\boldsymbol{C}=\alpha\boldsymbol{A}+\beta\boldsymbol{B} \tag{2-6}$$

$$c_{ij}=\alpha a_{ij}+\beta b_{ij}(i=1,\cdots,n;j=1,\cdots,m) \tag{2-7}$$

两个矩阵 $\boldsymbol{A}$ 和 $\mathbf{R}^{n\times k}$ 和 $\boldsymbol{B}\in\mathbf{R}^{k\times m}$ 的乘积为

$$\boldsymbol{C}=\boldsymbol{A}\boldsymbol{B} \tag{2-8}$$

应注意，矩阵 $\boldsymbol{A}$ 的列数必须等于矩阵 $\boldsymbol{B}$ 的行数。矩阵 $\boldsymbol{C}\in\mathbf{R}^{n\times m}$ 的元素为

$$c_{ij}=\sum_{l=1}^{k} a_{il}b_{lj} \tag{2-9}$$

两个相同维数的矩阵的 Hadamard 积是二者对应位置元素的乘积，即

$$\boldsymbol{C}=\boldsymbol{A}\odot\boldsymbol{B} \tag{2-10}$$

$$c_{ij}=a_{ij}b_{ij} \tag{2-11}$$

附注 2.1 在 MATLAB® 环境中，全矩阵乘法表示为 $\boldsymbol{C}=\boldsymbol{A}*\boldsymbol{B}$，Hadamard 积表示为 $\boldsymbol{C}=\boldsymbol{A}*\boldsymbol{B}$。

2.1.3 矩阵的逆

方阵 $\boldsymbol{A}$ 的逆矩阵 $\boldsymbol{A}^{-1}$ 必须满足

$$\boldsymbol{A}\boldsymbol{A}^{-1} = \boldsymbol{A}^{-1}\boldsymbol{A} = \boldsymbol{I} \tag{2-12}$$

式中:$\boldsymbol{I}$ 是与矩阵 $\boldsymbol{A}$ 有相同维度的单位矩阵。单位矩阵是主对角线上的元素为 1、其余元素均为 0 的方阵。并非所有矩阵都有逆矩阵。

没有逆矩阵的矩阵被称为奇异矩阵。逆矩阵的所有元素包含一个矩阵行列式的除数。判断一个矩阵是否是奇异矩阵,一个简单的方法是计算它的行列式。如果行列式是 0,那么矩阵是奇异的。

可以有多种方法计算逆矩阵。对于小型矩阵,可使用基于代数余子式和矩阵行列式的解析方法进行求解。而对于大矩阵,采用数值方法进行求解,例如 Gauss – Jordon 消元法或 LU 分解法。如果矩阵为对称正定矩阵,LU 分解法可简化为 Cholesky 分解法。具体技术细节可参考关于矩阵线性代数的文献。

附注 2.2 在 MATLAB® 环境中,有两个求解等式 $\boldsymbol{A} * \boldsymbol{x} = \boldsymbol{b}$ 的主要方法,其中一个是首选的。另一种不那么常用,但更简单的方法是直接使用 $\boldsymbol{x} = \boldsymbol{A}^{-1} * \boldsymbol{b}$:**x** = inv(A) * **b**。这里 inv(A) 是一个 MATLAB® 函数,解析计算 $\boldsymbol{A}^{-1}$,运算符 * 表示矩阵乘法。更好的解决方案是利用矩阵左除运算或者反斜杠:x = A\b。后者使用高效的 LU 分解法或者 Cholesky 分解法,这取决于矩阵 $\boldsymbol{A}$ 的性质。

需要注意的是,求逆矩阵时,可以用 B = A\eye(size(A)) 来高效求得,这里 $\boldsymbol{B}$ 是 $\boldsymbol{A}$ 的逆矩阵。eye(N) 是一个 MATLAB® 函数,可以用来得除了对角线元素为 1 外,其他元素都为 0 的单位矩阵。MATLAB® 函数 eye 还有其他的功能,例如可以生成非方矩阵,这里不再详细讨论。

如果矩阵线性代数方程是 $\boldsymbol{x} * \boldsymbol{A} = \boldsymbol{b}$,那么 $\boldsymbol{x} = \boldsymbol{b} * \boldsymbol{A}^{-1}$,和 MATLAB® 表达式 x = A/b 对矩阵进行表示。类似的,MATLAB 根据 $\boldsymbol{A}$ 和 $\boldsymbol{b}$ 的属性,选取最优的算法进行计算。

最后,如果只需要进行两个相同维数矩阵的对应位置元素的除法,可以用 MATLAB® 的 C = A./B 表达式。

2.1.4 分块矩阵的逆

对于分块矩阵,有一些求逆的通用公式。

引理 2.1 矩阵求逆引理:假设 $\boldsymbol{A}$、$\boldsymbol{C}$ 和 $(\boldsymbol{C}^{-1} + \boldsymbol{D}\boldsymbol{A}^{-1}\boldsymbol{B})$ 是非奇异方阵,那么

$$(\boldsymbol{A} + \boldsymbol{B}\boldsymbol{C}\boldsymbol{D})^{-1} = \boldsymbol{A}^{-1} - \boldsymbol{A}^{-1}\boldsymbol{B}(\boldsymbol{C}^{-1} + \boldsymbol{D}\boldsymbol{A}^{-1}\boldsymbol{B})^{-1}\boldsymbol{D}\boldsymbol{A}^{-1} \tag{2-13}$$

引理 2.2 分块矩阵求逆:将 $n \times m$ 矩阵 $\boldsymbol{M}$ 分块

$$\boldsymbol{M} = \begin{bmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{bmatrix} \begin{matrix} \}\, n \\ \}\, m \end{matrix} \quad \begin{matrix} \underbrace{}_{n} & \underbrace{}_{m} \end{matrix} \tag{2-14}$$

式中：$n \times n$ 矩阵 $\boldsymbol{A}$ 和 $m \times m$ 矩阵 $\boldsymbol{D}$ 是可逆的。两个矩阵的乘积为

$$\begin{bmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{bmatrix} \begin{bmatrix} \boldsymbol{X} & \boldsymbol{Y} \\ \boldsymbol{Z} & \boldsymbol{U} \end{bmatrix} = \begin{bmatrix} \boldsymbol{I}_n & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I}_m \end{bmatrix} \tag{2-15}$$

可以推知

$$\begin{bmatrix} \boldsymbol{X} & \boldsymbol{Y} \\ \boldsymbol{Z} & \boldsymbol{U} \end{bmatrix} \triangleq \begin{bmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{bmatrix}^{-1} \tag{2-16}$$

进而得到分块矩阵的逆为

$$\boldsymbol{M}^{-1} = \begin{bmatrix} (\boldsymbol{A} - \boldsymbol{B}\boldsymbol{D}^{-1}\boldsymbol{C})^{-1} & -\boldsymbol{A}^{-1}\boldsymbol{B}(\boldsymbol{D} + \boldsymbol{C}\boldsymbol{A}^{-1}\boldsymbol{B})^{-1} \\ -\boldsymbol{D}^{-1}\boldsymbol{C}(\boldsymbol{A} - \boldsymbol{B}\boldsymbol{D}^{-1}\boldsymbol{C})^{-1} & (\boldsymbol{D} + \boldsymbol{C}\boldsymbol{A}^{-1}\boldsymbol{B})^{-1} \end{bmatrix} \tag{2-17}$$

还可以通过下式计算分块矩阵的逆矩阵

$$\begin{bmatrix} \boldsymbol{X} & \boldsymbol{Y} \\ \boldsymbol{Z} & \boldsymbol{U} \end{bmatrix} \begin{bmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{bmatrix} = \begin{bmatrix} \boldsymbol{I}_n & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I}_m \end{bmatrix} \tag{2-18}$$

进而得相同分块矩阵的逆矩阵为

$$\boldsymbol{M}^{-1} = \begin{bmatrix} (\boldsymbol{A} - \boldsymbol{B}\boldsymbol{D}^{-1}\boldsymbol{C})^{-1} & -(\boldsymbol{A} - \boldsymbol{B}\boldsymbol{D}^{-1}\boldsymbol{C})^{-1}\boldsymbol{B}\boldsymbol{D}^{-1} \\ -(\boldsymbol{D} + \boldsymbol{C}\boldsymbol{A}^{-1}\boldsymbol{B})^{-1}\boldsymbol{C}\boldsymbol{A}^{-1} & (\boldsymbol{D} + \boldsymbol{C}\boldsymbol{A}^{-1}\boldsymbol{B})^{-1} \end{bmatrix} \tag{2-19}$$

2.1.5 矩阵的平方根

对任意非奇异方阵 $\boldsymbol{A}$，$\boldsymbol{A}$ 的平方根用矩阵 $\boldsymbol{B}$ 表示，满足等式

$$\boldsymbol{A} = \boldsymbol{B}\boldsymbol{B}^{\mathrm{T}} \tag{2-20}$$

它实际上不是平方根，而是一种矩阵分解。

附注2.3 在 MATLAB® 中有一个特殊的函数 B = sqrtm(A)，用来计算矩阵 $\boldsymbol{A}$ 的主(正)平方根。$\boldsymbol{B}$ 是每个特征值都有非负实部的唯一的平方根。如果 $\boldsymbol{A}$ 有非负实部的特征值，那么将得复数解。如果 $\boldsymbol{A}$ 是奇异的，那么 $\boldsymbol{A}$ 可能没有平方根。如果检测到确切的奇异点，MATLAB® 将给出出错信息。

另一种方法利用 Cholesky 分解法来计算下三角矩阵 $\boldsymbol{B}$，该矩阵有满足矩阵平方根性质的严格正对角线元素。$\boldsymbol{B}$ = chol($\boldsymbol{A}$) 由矩阵 $\boldsymbol{A}$ 生成一个上三角矩阵 $\boldsymbol{B}^{\mathrm{T}}$，满足等式 $\boldsymbol{A} = \boldsymbol{B}\boldsymbol{B}^{\mathrm{T}}$。假设下三角矩阵 $\boldsymbol{B}$ 是上三角矩阵 $\boldsymbol{B}^{\mathrm{T}}$ 的(复共轭)转置。矩阵 $\boldsymbol{A}$ 必须是正定的；否则 MATLAB® 将给出出错信息。

尽管 MATLAB® 中计算矩阵平方根的不同方法生成不同的结果矩阵 $\boldsymbol{B}$，但它们都是有效的，只是使用的数值技术不同。采用哪种方法取决于具体的应用。

2.2 矢量点发生器

在讨论 Sigma 点卡尔曼滤波器(第8～第12章)时，本书将利用一个简化计

数法来生成矢量 Sigma 点，进行多维数值积分。

定义 H 维矢量发生函数为

$$\boldsymbol{u} = (u_1, \cdots, u_r, 0, \cdots, 0) \in \mathbf{R}^n \tag{2-21}$$

其中，如果 $i \leqslant j$[2,3]，则 $0 \leqslant u_i \leqslant u_j$。在这种计数法中，$\mathbf{R}^n$ 指定 $\boldsymbol{u}$ 生成矢量点的维数，u_i 表示 $\pm u_i$。这种发生器可以表示为

$$[\boldsymbol{u}] \in \mathbf{R}^n \tag{2-22}$$

或者

$$[u_1, \cdots, u_r] \in \mathbf{R}^n \tag{2-23}$$

式中：$\boldsymbol{u} \triangleq [u_1, \cdots, u_r]$ 为方便表达，隐藏了零坐标。这种计数法表示的 k 维矢量点集合，可以从对 $\boldsymbol{u}$ 进行置换或交换若干坐标记号来产生。

在第 Ⅱ 部分，将研究高斯加权积分的多维积分规则。数值解法涉及矢量点集非线性函数的评价，包括一个常数乘以单位超球面或超立方体上的矢量点集合。因此，本书会利用发生器来生成形式的矢量点集合。注意 $[1,1,1] \in \mathbf{R}^n$ 与 $[\pm 1, \pm 1, \pm 1, 0, \cdots, 0]$ 所有可能的排列是相同的，其中有 $n-3$ 个 0。下面给出一些实例。在这些例子中，使用符号[3]

$$n^{(k)} \triangleq n(n-1)\cdots(n-k+1) \tag{2-24}$$

$$n_{(k)} \triangleq \frac{n(n+1)\cdots(n-k+1)}{k!} \tag{2-25}$$

考虑 $u = 1$ 的情况。下面例子给出了各种情况下生成的矢量点。

例 2.1　对于 $[1] \in \mathbf{R}^2$，矢量点集合由矢量组成

$$[1] \in \mathbf{R}^2 \triangleq \begin{cases} [1,0]^{\mathrm{T}} \\ [0,1]^{\mathrm{T}} \\ [-1,0]^{\mathrm{T}} \\ [0,-1]^{\mathrm{T}} \end{cases} \tag{2-26}$$

对于 $\mathbf{R}^2$ 空间，当 $n=2$ 时，在笛卡儿空间的二维坐标轴上会有 $2n=4$ 个矢量点，如图 2-1 所示。

例 2.2　对于 $[1] \in \mathbf{R}^3$，矢量点集包含

$$[1] \in \mathbf{R}^3 \triangleq \begin{cases} [1,0,0]^{\mathrm{T}} \\ [0,1,0]^{\mathrm{T}} \\ [0,0,1]^{\mathrm{T}} \\ [-1,0,0]^{\mathrm{T}} \\ [0,-1,0]^{\mathrm{T}} \\ [0,0,-1]^{\mathrm{T}} \end{cases} \tag{2-27}$$

这里，矢量点落到三维超球面的笛卡儿坐标轴上，如图 2-2 所示。

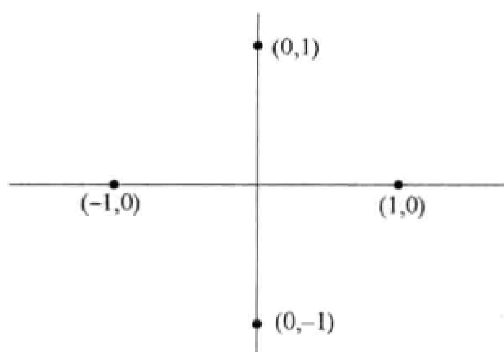


图 2 - 1　二维笛卡儿网格坐标轴矢量点示例

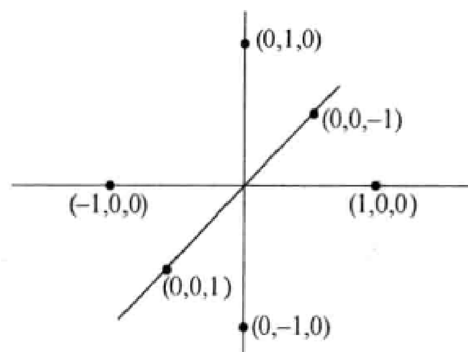


图 2 - 2　三维笛卡儿网格坐标轴矢量点示例

例 2.3　对于$[1,1] \in \mathbf{R}^2$,矢量点集合包含$2n^{(2)} = 2n(n-1) = 4$个矢量点,即

$$[1,1] \in \mathbf{R}^2 \triangleq \begin{cases} [1,1]^{\mathrm{T}} \\ [1,-1]^{\mathrm{T}} \\ [-1,1]^{\mathrm{T}} \\ [-1,-1]^{\mathrm{T}} \end{cases} \tag{2-28}$$

例 2.4　对于$[1,1] \in \mathbf{R}^3$,集合中有$2n^{(2)} = 2n(n-1)(n-2) = 2 \cdot 3 \cdot 2 \cdot 1 = 12$个矢量点,即

$$[1,1] \in \mathbf{R}^3 \triangleq \begin{cases} [1,1,0]^{\mathrm{T}} \\ [1,0,1]^{\mathrm{T}} \\ [0,1,1]^{\mathrm{T}} \\ [-1,1,0]^{\mathrm{T}} \\ [-1,0,1]^{\mathrm{T}} \\ [0,-1,1]^{\mathrm{T}} \\ [1,-1,0]^{\mathrm{T}} \\ [1,0,-1]^{\mathrm{T}} \\ [0,1,-1]^{\mathrm{T}} \\ [-1,-1,0]^{\mathrm{T}} \\ [-1,0,-1]^{\mathrm{T}} \\ [0,-1,-1]^{\mathrm{T}} \end{cases} \tag{2-29}$$

通过式(2 - 24)和式(2 - 25)可知,由一个发生器产生的矢量点的数目可以由下式[3]得,即

$$
\begin{array}{cc}
\text{发生器} & \text{矢量点数目} \\
[0] \in \mathbf{R}^n & 1 \\
[1] \in \mathbf{R}^n & 2n \\
[1,1] \in \mathbf{R}^n & 2n_{(2)} \\
[1,1,1] \in \mathbf{R}^n & 2^3 n_{(3)} \\
[1,1,1,1] \in \mathbf{R}^n & 2^4 n_{(4)} \\
\vdots & \vdots
\end{array}
\tag{2-30}
$$

2.3 多参数多维非线性函数估计

对于许多跟踪应用,估计方法需要用一种积分值进行评价,这种积分值是多维参数乘以一个概率密度得的非线性多维函数的积分。本节将回顾数值逼近非线性函数方法。

本节首先分析标量方法,然后扩展到多维情况。所有非线性函数的近似实质上都是将非线性函数扩展为一个具有任意系数的多项式。系数的具体取值取决于实际应用,根据精度要求和其他计算考虑因素来决定。估计和跟踪应用中,系数的取值规范将在后面关于研究特定方法进行密度加权积分评价的章节中予以解决。

2.3.1 节回顾一下用标量化参数逼近标量非线性函数的方法。首先讨论标量函数的一般多项式展开,在此基础上推导出非线性函数的标量泰勒近似多项式。用有限差分替代泰勒多项式的差值,导出泰勒多项式的数值逼近,得 Stirling 插值公式(Stirling 多项式)。

接下来的小节中,通过多维多项式扩展的介绍,将标量估计推广为具有多维参数的多维非线性函数的估计。这一推导再次得广义多维泰勒多项式。将 Stirling 逼近推广至多个维度可以通过多种方式实现,其中,每一种都会引出一个不同的数值估计方法。要导出一个简单的多维 Stirling 估计,需要 3^{n_x} 个笛卡儿插值点,其中 n_x 是多维函数参数的维度。下面简要讨论两种简化方法。第一种简化方法中,插值点可以减少至只包含那些落在坐标轴上的点,数目减少至 $2n_x+1$;第二种简化方法是使用单形 Stirling 逼近,将插值点数目减少至 n_x+2。

本书的讨论内容仅限于估计方法,几乎不会涉及同样也很重要的逼近误差估计和收敛性等内容。

2.3.1 标量非线性函数估计

2.3.1.1 一般多项式展开

通常来说,任意一个参数为标量的连续非线性标量化函数 $f(x)$,都可以用

任意点 x_0,通过一系列的逼近得任意精度的估计,即

$$f(x) = \sum_{i=0}^{\infty} a_i(x - x_0)^i = a_0 + a_1(x - x_0) + a_2(x - x_0)^2 + a_3(x - x_0)^3 + \cdots \quad (2-31)$$

通常很难找到一个在无穷范围求和的封闭形式的解 ∞,因此将序列截断为一个 M 阶的多项式估计,即

$$f(x) \simeq \sum_{i=0}^{M} a_i(x - x_0)^i = a_0 + a_1(x - x_0) + a_2(x - x_0)^2 + \cdots + a_M(x - x_0)^M \quad (2-32)$$

式中:$a_i(x - x_0)^i$ 这个多项式的每一项都可以进一步展开为带有系数 b_i 的单项式 x^i。因此,函数 $f(x)$ 的多项式估计可以写成一个展开式,即

$$f(x) \simeq \sum_{i=0}^{M} b_i x^i = b_0 + b_1 x + b_2 x^2 + \cdots + b_M x^M \quad (2-33)$$

$$b_0 \triangleq \sum_{i=0}^{M} (-1)^i a_i x_0^i$$

$$b_1 \triangleq \sum_{i=1}^{M} (-2)^{i-1} a_i x_0^{i-1}$$

$$\vdots$$

$$b_M \triangleq a_M \quad (2-34)$$

2.3.1.2 泰勒多项式展开

下面给出 M 阶泰勒多项式和麦克劳林多项式的简单推导过程。前提条件是 $f(x)$ 的 M 阶导数存在。令式(2-32)中的 $x = x_0$,得

$$a_0 = [f(x)]_{x=x_0} \quad (2-35)$$

为求解 a_1,对式(2-32)求解关于变量 x 的导数,并设 $x = x_0$,得

$$a_1 = \left[\frac{\mathrm{d}}{\mathrm{d}x} f(x)\right]_{x=x_0} \quad (2-36)$$

求式(2-32)在 $x = x_0$ 处的各级导数,对于每种情况,用 $f(x)$ 的高阶导数对所有系数进行评价,导出著名的 $f(x_0)$ 的标量化泰勒多项式点逼近,即

$$f(x_0) \simeq [f(x)]_{x=x_0} + \left[\frac{\mathrm{d}}{\mathrm{d}x} f(x)\right]_{x=x_0}(x - x_0) + \cdots + \frac{1}{M!}\left[\left(\frac{\mathrm{d}}{\mathrm{d}x}\right)^M f(x)\right]_{x=x_0}(x - x_0)^M \quad (2-37)$$

$$= \sum_{i=0}^{M} \frac{1}{i!}(x - x_0)^i \left[\left(\frac{\mathrm{d}}{\mathrm{d}x}\right)^i f(x)\right]_{x=x_0} \quad (2-38)$$

通过对 $f(x)$ 的运算定义标量算子如下

$$D_x^i f(x) \triangleq (x - x_0)^i \left[\left(\frac{\mathrm{d}}{\mathrm{d}x} \right)^i f(x) \right]_{x = x_0} \tag{2-39}$$

M 阶泰勒多项式(2 - 38) 变换为

$$f(x) = \sum_{i=0}^{M} \frac{1}{i!} D_x^i f(x) \tag{2-40}$$

有些情况下,计算 $f(x)$ 的高阶导数非常困难,泰勒多项式逼近通常在不超过三项($M = 2$) 的时候被截断。在 Arfken[4] 中可以找到一个略微不同的泰勒多项式推导,得出由于截断泰勒多项式产生的误差。M 阶泰勒多项式的误差项如下

$$R_M = \frac{(x - x_0)^M}{N!} \left(\frac{\mathrm{d}}{\mathrm{d}\xi} \right)^M f(\xi) \tag{2-41}$$

式中:$x_0 \leqslant \xi \leqslant x$。

2.3.1.3 Stirling 多项式展开

还有一种 $f(x)$ 无导数逼近方法,在泰勒多项式逼近中利用中心差分法来代替求导数。通过泰勒级数展开式得有限差分逼近。假设函数 $f(x)$ 在感兴趣区域连续可微,并且 $f(x_0)$ 及其在 $x = x_0$ 处的所有导数已知。根据式(2 - 37),检验 $f(x_0 + q)$ 关于 x_0 的一阶泰勒多项式

$$f(x_0 + q) = f(x_0) + f'(x_0)q \tag{2-42}$$

$$q \triangleq x - x_0$$

$$f'(x_0) \triangleq \mathrm{d}f(x_0)/\mathrm{d}x_0$$

按照相同的步骤,得

$$f(x_0 - q) = f(x_0) - f'(x_0)q \tag{2-43}$$

将两个等式相减,得一阶中心差分逼近

$$f'(x) = \frac{1}{2q}[f(x + q) - f(x - q)] \tag{2-44}$$

这里用 x_i 代替 x_0。

对于二阶泰勒多项式,按照同样步骤计算,得

$$f(x + q) = f(x) + f'(x)q + \frac{1}{2}f''(x)q^2 \tag{2-45}$$

$$f(x - q) = f(x) - f'(x)q + \frac{1}{2}f''(x)q^2 \tag{2-46}$$

$$f''(x) \triangleq \mathrm{d}^2 f(x_0)/\mathrm{d}x^2$$

将上面两个等式相加,得

$$f''(x) = \frac{1}{q^2}[f(x + q) - 2f(x) + f(x - q)] \tag{2-47}$$

利用上述导数项中心差分逼近,将二阶泰勒多项式(2 - 37) 简化为二阶 Stirling 插值公式[5],即

$$f(x) = f(x_0) + \tilde{D}_x f(x) + \frac{1}{2}\tilde{D}_x^2 f(x) \tag{2-48}$$

$$\widetilde{D}_x f(x) \triangleq \frac{1}{2q}[f(x_0+q)-f(x_0-q)](x-x_0) \tag{2-49}$$

$$\widetilde{D}_x^2 f(x) \triangleq \frac{1}{q^2}[f(x_0+q)-2f(x_0)+f(x_0-q)](x-x_0)^2 \tag{2-50}$$

当 $x_0=0$ 时,逼近结果显而易见。

上面就是对$f(x)$无导数逼近的方法,步长参数q可取任意值。对上面二阶逼近算法进行检验可知,在一维情况下,$f(x)$在$\{x_0,x_0+q,x_0-q\}$这三点必须已知。高阶逼近可以通过高阶导数近似来简单地得,得间距为q的一维均匀网格上评估点数目的增加。

为总结标量方法,本书给出了函数$f(x)$的两种多项式逼近方法:泰勒多项式扩展和Stirling插值公式扩展。针对上述两种情况,本书检验了点逼近,即$f(x)$在点处的多项式逼近x_0。大支撑域非线性函数的多项式逼近则是一个更广泛的主题,超出了本书的研究范畴。

2.3.2 多维非线性函数估计

2.3.2.1 一般多维多项式展开

对于多维情况,可以将非线性函数写成矢量参数$\boldsymbol{f}(\boldsymbol{x})$,其中且$\boldsymbol{f}\in\mathbf{R}^{n_f}$且$\boldsymbol{x}\in\mathbf{R}^{n_x}$

$$\boldsymbol{f}(\boldsymbol{x})=[f_1(\boldsymbol{x}),f_2(\boldsymbol{x}),\cdots,f_{n_f}(\boldsymbol{x})]^{\mathrm{T}} \tag{2-51}$$

$$\boldsymbol{x}=[x_1,x_2,\cdots,x_{n_x}]^{\mathrm{T}} \tag{2-52}$$

注意矢量函数$\boldsymbol{f}$可能和$\boldsymbol{x}$维数不同。对于参数为矢量的标量化非线性函数,一般的关于点$\boldsymbol{x}_0\triangleq[x_{0,1},x_{0,2},\cdots,x_{0,n_x}]^{\mathrm{T}}$的$d$阶多项式展开可以写成多维单项式的形式[2],即

$$\begin{aligned} f(\boldsymbol{x}) &= f(x_1,x_2,\cdots,x_{n_x}) \\ &\simeq \sum_{i_1=0}^{d}\cdots\sum_{i_{n_x}=0}^{d} a_j(i_1,i_2,\cdots,i_{n_x})\cdot \\ &\quad (x_1-x_{0,1})^{i_1}(x_2-x_{0,2})^{i_2}\cdots(x_{n_x}-x_{0,n_x})^{i_{n_x}} \end{aligned} \tag{2-53}$$

其中$i_1+i_2+\cdots+i_{n_x}\leqslant d$表明只有满足特定条件的多重和项能进行多项式展开。从下例中可以清晰地看到这一点。

例 2.5 对于$\boldsymbol{x}\in\mathbf{R}^2$,由$d=2$或更低阶次基底张成的多维单项式由一组6个单项式构成$\{1,(x_1-x_{0,1}),(x_2-x_{0,2}),(x_1-x_{0,1})(x_2-x_{0,2}),(x_2-x_{0,1})^2,(x_2-x_{0,2})^2\}$。$f(\boldsymbol{x})$关于点$x_0$的多项式逼近如下所示

$$\begin{aligned} f(\boldsymbol{x}) = f(x_1,x_2) &\simeq a(0,0)+ \\ &a(1,0)(x_1-x_{0,1})+a(0,1)(x_2-x_{0,2})+ \\ &a(1,1)(x_1-x_{0,1})(x_2-x_{0,2})+ \end{aligned}$$

$$a(2,0)(x_1 - x_{0,1})^2 + a(0,2)(x_2 - x_{0,2})^2 \tag{2-54}$$

按照式(2 - 53) 的形式,可写成

$$f(\boldsymbol{x}) = \sum_{i_1=0}^{d}\sum_{i_2=0}^{d} a(i_1,i_2)(x_1 - x_{0,1})^{i_1}(x_2 - x_{0,2})^{i_2} \tag{2-55}$$

其中多项式中的每一个单项式必须满足条件 $i_1 + i_2 \leqslant d = 2$。

一般来说,对于矢量 $\boldsymbol{x} \in \mathbf{R}^{n_x}$ 有 $\binom{n_x + d}{d}$ 个精度为 d 或者更少的单项式。二项式系数 $\binom{n}{m}$ 是从 n 个不同对象的集合中选取 m 个对象的组合数(不考虑顺序),定义为

$$\binom{n}{m} \triangleq \frac{n!}{m!(n-m)!} \tag{2-56}$$

对于矢量函数 $\boldsymbol{f}(\boldsymbol{x})$ 关于点的多项式展开 $\boldsymbol{x}_0$ 可写成

$$\begin{aligned}\boldsymbol{f}(\boldsymbol{x}) &= \boldsymbol{f}(x_1, x_2, \cdots, x_{n_x}) \\ &\simeq \sum_{i_1=0}^{d}\cdots\sum_{i_{n_x}=0}^{d}\boldsymbol{a}(i_1, i_2, \cdots, i_{n_x}) \cdot \\ &\quad (x_1 - x_{0,1})^{i_1}(x_2 - x_{0,2})^{i_2}\cdots(x_{n_x} - x_{0,n_x})^{i_{n_x}}\end{aligned} \tag{2-57}$$

$$\boldsymbol{a} = [a_1, a_2, \cdots, a_{n_f}]^{\mathrm{T}} \tag{2-58}$$

例 2.6 下面再次考虑 $x \in \mathbf{R}^2$ 且最高为 2 次幂的情况,式(2 - 57) 变换为

$$\begin{aligned}\boldsymbol{f}(\boldsymbol{x}) \approx\ & \boldsymbol{a}(0,0) + \boldsymbol{a}(1,0)(x_1 - x_{0,1}) + \boldsymbol{a}(0,1)(x_2 - x_{0,2}) + \\ & \boldsymbol{a}(1,1)(x_1 - x_{0,1})(x_2 - x_{0,2}) + \\ & \boldsymbol{a}(2,0)(x_1 - x_{0,1})^2 + \boldsymbol{a}(0,2)(x_2 - x_{0,2})^2\end{aligned} \tag{2-59}$$

后面两节中将首先解决多维函数的泰勒多项式逼近问题,然后讨论将 Stirling 插值公式向多维推广的问题。

2.3.2.2 多维泰勒多项式展开

在 $\boldsymbol{x} \in \mathbf{R}^2$ 且 $\boldsymbol{f}(\boldsymbol{x}) \in \mathbf{R}^{n_f}$ 的情况下,假定式(2 - 59) 中的 $\boldsymbol{x} = \boldsymbol{x}_0$,得泰勒多项式

$$\boldsymbol{a}(0,0) = [\boldsymbol{f}(\boldsymbol{x})]_{x=x_0} \tag{2-60}$$

然后,对式(2 - 59) 中 $\boldsymbol{f}(\boldsymbol{x})$ 关于 x_1 取偏导数,令 $\boldsymbol{x} = \boldsymbol{x}_0$ 可导出

$$\boldsymbol{a}(1,0) = \left[\frac{\partial}{\partial x_1}\boldsymbol{f}(\boldsymbol{x})\right]_{x=x_0} \tag{2-61}$$

$\partial/\partial x_1$ 对 $\boldsymbol{f}(\boldsymbol{x})$ 中每个元素执行运算。

用类似方法,由式(2 - 59) 确定其余系数的取值,得如下的二阶多项式展开,即

$$f(\boldsymbol{x}) \simeq [f(\boldsymbol{x})]_{\boldsymbol{x}=\boldsymbol{x}_0} + \left\{(x_1 - x_{0,1})\left[\frac{\partial}{\partial x_1}f(\boldsymbol{x})\right]_{\boldsymbol{x}=\boldsymbol{x}_0} + (x_2 - x_{0,2})\left[\frac{\partial}{\partial x_2}f(\boldsymbol{x})\right]_{\boldsymbol{x}=\boldsymbol{x}_0}\right\} + \frac{1}{2!}\left\{(x_1 - x_{0,1})^2\left[\frac{\partial^2}{\partial x_1^2}f(\boldsymbol{x})\right]_{\boldsymbol{x}=\boldsymbol{x}_0} + 2(x_1 - x_{0,1})(x_2 - x_{0,2})\left[\frac{\partial^2}{\partial x_1 \partial x_2}f(\boldsymbol{x})\right]_{\boldsymbol{x}=\boldsymbol{x}_0} + (x_2 - x_{0,2})^2\left[\frac{\partial^2}{\partial x_2^2}f(\boldsymbol{x})\right]_{\boldsymbol{x}=\boldsymbol{x}_0}\right\} \tag{2-62}$$

这可以很容易推广到 M 次多维泰勒多项式逼近，即

$$f(\boldsymbol{x}) \simeq \sum_{i=0}^{M} \frac{1}{i!} D_{\boldsymbol{x}-\boldsymbol{x}_0}^{i} f(\boldsymbol{x}) \tag{2-63}$$

运算符号类似于 Williamson 等[6] 所采用的运算符，用下式定义标量运算符 $D_{\boldsymbol{x}-\boldsymbol{x}_0}^{i}$，即

$$D_{\boldsymbol{x}-\boldsymbol{x}_0}^{i} f(\boldsymbol{x}) \triangleq \left[\frac{\partial}{\partial x_1}(x_1 - x_{0,1}) + \frac{\partial}{\partial x_2}(x_2 - x_{0,2}) + \cdots + \frac{\partial}{\partial x_{n_x}}(x_{n_x} - x_{0,n_x})\right]^{i} f(\boldsymbol{x})\bigg|_{\boldsymbol{x}=\boldsymbol{x}_0} \tag{2-64}$$

符号 $[\cdot]|_{\boldsymbol{x}=\boldsymbol{x}_0}$ 仅适用于微分项。将多项式 $(x_1 + \cdots + x_n)^i$ 展开，每一项都由 $x_1^{i_1}x_2^{i_2}\cdots x_n^{i_n}$ 形式的因子的常数倍构成，其中非负整数 i_k 满足 $i_1 + \cdots + i_n = i$。多项式展开形式如下

$$(x_1 + \cdots + x_n)^i = \sum_{\substack{i_1=0 \\ i_1+\cdots+i_n=i}}^{i} \cdots \sum_{i_n=0}^{i} \binom{i}{i_1 \cdots i_n} x_1^{i_1} \cdots x_n^{i_n} \tag{2-65}$$

多项式系数可由下式求得，即

$$\binom{i}{i_1 \cdots i_n} \triangleq \frac{i!}{i_1! \cdots i_n!} \tag{2-66}$$

例 2.7 考虑这种情况

$$(x_1 + x_2 + x_3)^2 = \sum_{\substack{i_1=0 \\ i_1+i_2+i_3=2}}^{2} \sum_{i_2=0}^{2} \sum_{i_3=0}^{2} \binom{2}{i_1 \cdots i_3} x_1^{i_1} x_2^{i_2} x_3^{i_3} \tag{2-67}$$

$$= x_1^2 + x_2^2 + x_3^2 + 2x_1x_2 + 2x_1x_3 + 2x_2x_3 \tag{2-68}$$

式(2－64)可以直接写成

$$D_{\boldsymbol{x}-\boldsymbol{x}_0}^{i} f(\boldsymbol{x}) = \sum_{\substack{i_1=0 \\ i_1+\cdots+i_n=i}}^{i} \cdots \sum_{i_n=0}^{i} \binom{i}{i_1 \cdots i_n} (x_1 - x_{0,1})^{i_1} \cdots (x_{n_x} - x_{0,n_x})^{i_n} \cdot \left[\frac{\partial^i}{\partial x_1^{i_1} \cdots \partial x_{n_x}^{i_n}} f(\boldsymbol{x})\right]_{\boldsymbol{x}=\boldsymbol{x}_0} \tag{2-69}$$

将式(2－63)转换为用矢量矩阵表示，只保留二阶以下的项，泰勒多项式逼近结果为一个多维函数

$$f(\boldsymbol{x}) = [f(\boldsymbol{x})]_{\boldsymbol{x}=\boldsymbol{x}_0} + [\nabla_{\boldsymbol{x}} \boldsymbol{f}^{\mathrm{T}}(\boldsymbol{x})]^{\mathrm{T}}_{\boldsymbol{x}=\boldsymbol{x}_0}(\boldsymbol{x} - \boldsymbol{x}_0) + \frac{1}{2!}\sum_{p=1}^{n_x} \boldsymbol{e}_p(\boldsymbol{x} - \boldsymbol{x}_0)^{\mathrm{T}}[\nabla_{\boldsymbol{x}} \nabla_{\boldsymbol{x}}^{\mathrm{T}} f_p(\boldsymbol{x})]_{\boldsymbol{x}=\boldsymbol{x}_0}(\boldsymbol{x} - \boldsymbol{x}_0) \tag{2-70}$$

一阶项是雅可比矩阵$[\nabla_{\boldsymbol{x}} \boldsymbol{f}^{\mathrm{T}}(\boldsymbol{x})]^{\mathrm{T}}_{\boldsymbol{x}=\boldsymbol{x}_0} \in \mathbf{R}^{n_x \times n_f}$，二阶项是一个 Hessian 矩阵，第 p 个分量为$[\nabla_{\boldsymbol{x}} \nabla_{\boldsymbol{x}}^{\mathrm{T}} f_p(\boldsymbol{x})]_{\boldsymbol{x}=\boldsymbol{x}_0} \in \mathbf{R}^{n_x \times n_f}$，$\boldsymbol{e}_p$ 是一个沿 p 维笛卡儿坐标轴方向的 n_x 维单位矢量。很难计算多维函数的超出一阶项的泰勒多项式展开，所以通常截断多项式序列，使其没有高阶项。

2.3.2.3 多维 Stirling 多项式展开

多维函数无导数 Stirling 多项式插值公式可写成[2,7,8]，即

$$f(\boldsymbol{x}) \simeq \sum_{i=0}^{M} \frac{1}{i!} \tilde{D}^{i}_{\boldsymbol{x}-\boldsymbol{x}_0} f(\boldsymbol{x}) \tag{2-71}$$

或者仅保留至二阶为

$$f(\boldsymbol{x}) \simeq f(\boldsymbol{x}_0) + \tilde{D}_{\boldsymbol{x}-\boldsymbol{x}_0} f(\boldsymbol{x}) + \frac{1}{2!}\tilde{D}^{2}_{\boldsymbol{x}-\boldsymbol{x}_0} f(\boldsymbol{x}) \tag{2-72}$$

中心差分算子 $\tilde{D}_{\boldsymbol{x}-\boldsymbol{x}_0} f(\boldsymbol{x})$ 和 $\tilde{D}^{2}_{\boldsymbol{x}-\boldsymbol{x}_0} f(\boldsymbol{x})$ 定义为

$$\tilde{D}_{\boldsymbol{x}-\boldsymbol{x}_0} f(\boldsymbol{x}) \triangleq \frac{1}{2q}\sum_{p=1}^{n_x}(x_p - x_{0,p})[f(\boldsymbol{x}_0 + q\boldsymbol{e}_p) - f(\boldsymbol{x}_0 - q\boldsymbol{e}_p)] \tag{2-73}$$

$$\begin{aligned}\tilde{D}^{2}_{\boldsymbol{x}-\boldsymbol{x}_0} f(\boldsymbol{x}) \triangleq \frac{1}{q^2}\Big\{ &\sum_{p=1}^{n_x}(x_p - x_{0,p})^2[f(\boldsymbol{x}_0 + q\boldsymbol{e}_p) - 2f(\boldsymbol{x}_0) + f(\boldsymbol{x}_0 - q\boldsymbol{e}_p)] + \\ &2\sum_{p=1}^{n_x}\sum_{\substack{r=1\\ r\neq p}}^{n_x}(x_p - x_{0,p})(x_r - x_{0,r}) \cdot \\ &[f(\boldsymbol{x}_0 + q\boldsymbol{e}_p + q\boldsymbol{e}_r) - f(\boldsymbol{x}_0 - q\boldsymbol{e}_p + q\boldsymbol{e}_r) - \\ &f(\boldsymbol{x}_0 + q\boldsymbol{e}_p - q\boldsymbol{e}_r) + f(\boldsymbol{x}_0 - q\boldsymbol{e}_p - q\boldsymbol{e}_r)]\Big\}\end{aligned} \tag{2-74}$$

式中：$\boldsymbol{e}_p$ 和 $\boldsymbol{e}_r$ 分别为沿着第 p 维和第 r 维笛卡儿坐标轴的单位矢量。

例 2.8 对于 $\boldsymbol{x} \in \mathbf{R}^2$，得

$$\begin{aligned}\tilde{D}_{\boldsymbol{x}-\boldsymbol{x}_0} f(\boldsymbol{x}) \triangleq \frac{1}{2q}\{&(x_1 - x_{0,1})[f(\boldsymbol{x}_0 + q\boldsymbol{e}_p) - f(\boldsymbol{x}_0 - q\boldsymbol{e}_1)] + \\ &(x_2 - x_{0,2})[f(\boldsymbol{x}_0 + q\boldsymbol{e}_2) - f(\boldsymbol{x}_0 - q\boldsymbol{e}_2)]\}\end{aligned} \tag{2-75}$$

$$\begin{aligned}\tilde{D}^{2}_{\boldsymbol{x}-\boldsymbol{x}_0} f(\boldsymbol{x}) \triangleq &\frac{1}{q^2}(x_1 - x_{0,1})^2[f(\boldsymbol{x}_0 + q\boldsymbol{e}_1) - 2f(\boldsymbol{x}_0) + f(\boldsymbol{x}_0 - q\boldsymbol{e}_1)] + \\ &\frac{1}{q^2}(x_2 - x_{0,2})^2[f(\boldsymbol{x}_0 + q\boldsymbol{e}_2) - 2f(\boldsymbol{x}_0) + f(\boldsymbol{x}_0 - q\boldsymbol{e}_2)] +\end{aligned}$$

$$\frac{2}{q^2}(x_1 - x_{0,1})(x_2 - x_{0,2}) \cdot$$
$$[f(\boldsymbol{x}_0 + q\boldsymbol{e}_1 + q\boldsymbol{e}_2) - f(\boldsymbol{x}_0 - q\boldsymbol{e}_1 + q\boldsymbol{e}_2) -$$
$$f(\boldsymbol{x}_0 + q\boldsymbol{e}_1 - q\boldsymbol{e}_2) + f(\boldsymbol{x}_0 - q\boldsymbol{e}_1 - q\boldsymbol{e}_2)] \quad (2-76)$$

显然,插值 $\boldsymbol{x} \in \mathbf{R}^2$ 需要一组均匀分布在二维网格上的 9 个间距为 q 的点,$J_2 \in \{(x_{0,1}, x_{0,2}), (x_{0,1} \pm q, x_{0,2}), (x_{0,1}, x_{0,2} \pm q), (x_{0,1} \pm q, x_{0,2} \pm q)\}$,如图2－3所示。

通常,二阶多维 Stirling 多项式逼近所需点的个数为 3^{n_x},在均匀多维网格(立方体)上,间距为 q。

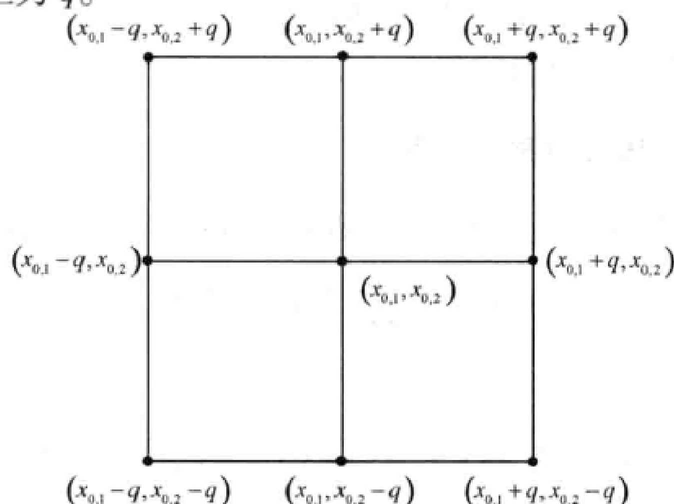


图 2－3　多维 Sterling 逼近的二维点选取

另外一种方法利用 Hermite 正交多项式[5,9] 函数展开式来代替等式(2－63)和式(2－69),得相同数目的插值点作为二阶多维 Stirling 逼近(插值点可能并不相同)。这种方法不需要进行求导或者进行有限差分运算,在很多方面优于有限差分方法。当然,以上陈述纯属个人的主观意见。这种插值法将在后续章节加以介绍。

之所以把 Hermite 插值内容介绍放在后面,是因为可以通过一个简单的变换,把高斯加权积分转化成 Hermite 积分,使算法更易于理解。

还有其他的无导数插值方法,比上述多维 Stirling 多项式逼近方法需要更少的插值点。这种插值方法只需要用到图2－3中那些落在正交笛卡儿坐标轴以及原点上的插值点。对比式(2－73)和式(2－74),可发现式(2－74)中只有交叉项被去掉了。对于二阶项来说,精度损失的代价非常小。经过这一简化,插值点的数目减少至 $2n_x + 1$。

另一种多维插值方法利用原点上的点和另外的多维超球面上等角间距点进行插值,这种方法称为球面单形插值。图2－4给出一个顶点在超球面上的三维正则单形图(所有边和角都相等)的例子。这些方法仅需要 $n_x + 2$ 个点即可实现二阶插值。虽然对于多维函数来说,单形插值方法是数值积分方法的基础,但直到最近,采用单形方法进行多维插值才见诸文献[10－16],具体为计算机图形

学相关应用。因此,单形插值方法的详细介绍超出了本书的范围。尽管如此,本书将在第 10 章给出用单形法进行多维函数的高斯加权积分。

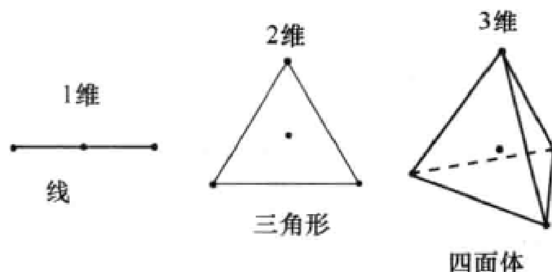


图 2－4　1－3 维单形图

2.4　多变量统计概述

2.4.1　一般定义

从贝叶斯观点来看,随机变量是指变量的值无法预先确定,可以根据概率分布改变的量。随机变量通常是多维的,所以其概率分布是多元的。此外,随机变量的期望值可以随时间变化,分布随时间的变化由动态方程或转移方程决定。

2.4.1.1　概率密度函数

这一节首先讨论概率密度函数和多维随机变量的累积分布函数的一般定义。

多维随机变量 $\boldsymbol{x}$ 在点 $\boldsymbol{x}=\boldsymbol{\xi}$ 处的概率密度函数可定义为[17]

$$p_{\boldsymbol{x}}(\boldsymbol{\xi})\mathrm{d}\boldsymbol{\xi} \triangleq \mathrm{Prob}\{\boldsymbol{\xi} \leqslant \boldsymbol{x} \leqslant \boldsymbol{\xi}+\mathrm{d}\boldsymbol{\xi}\} \triangleq P\{\boldsymbol{\xi} \leqslant \boldsymbol{x} \leqslant \boldsymbol{\xi}+\mathrm{d}\boldsymbol{\xi}\} > 0 \tag{2-77}$$

其中,$\mathrm{d}\boldsymbol{\xi}$ 是一个无限小区间。得出 $\boldsymbol{x}$ 取值在区间 $\boldsymbol{\eta} \leqslant \boldsymbol{x} \leqslant \boldsymbol{\xi}$ 内的概率为

$$\mathrm{Prob}\{\boldsymbol{\eta} \leqslant \boldsymbol{x} \leqslant \boldsymbol{\xi}\} \triangleq P\{\boldsymbol{\eta} \leqslant \boldsymbol{x} \leqslant \boldsymbol{\xi}\} = \int_{\substack{\boldsymbol{\eta} \\ \mathbf{R}^{n_x}}}^{\boldsymbol{\xi}} p_{\boldsymbol{x}}(\boldsymbol{x})\mathrm{d}\boldsymbol{x} \tag{2-78}$$

其中 n_x 为 $\boldsymbol{x}$ 的维度。多维积分定义如下

$$\begin{aligned}\int_{\substack{\boldsymbol{\eta} \\ \mathbf{R}^{n_x}}}^{\boldsymbol{\xi}} p_{\boldsymbol{x}}(\boldsymbol{x})\mathrm{d}\boldsymbol{x} &\triangleq \int_{\eta_1}^{\xi_1}\mathrm{d}x_1\cdots\int_{\eta_{n_x}}^{\xi_{n_x}}\mathrm{d}x_{n_x}p_{\boldsymbol{x}}(x_1,\cdots,x_{n_x}) \\ &= \prod_{i=1}^{n_x}\int_{\eta_i}^{\xi_i}\mathrm{d}x_i p_{\boldsymbol{x}}(p_1,\cdots,x_{n_x})\end{aligned} \tag{2-79}$$

$\boldsymbol{x}$ 落在区间 $\boldsymbol{\eta} \leqslant \boldsymbol{x} \leqslant \boldsymbol{\xi}$ 的概率为区间中概率密度函数的多维“体积”。概率密度函数更常用的表达式是用 $p(\boldsymbol{x})$ 来表示 $p_{\boldsymbol{x}}(\boldsymbol{x})$。

2.4.1.2　累积分布函数

累积分布函数定义为

$$P_x\{\xi\} \triangleq \text{Prob}\{x \leqslant \xi\} = \int_{-\infty \atop \mathbf{R}^{n_x}}^{\xi} p_x(x)\,\mathrm{d}x \tag{2-80}$$

注意当$\boldsymbol{\xi} \to \infty$ 时,由于概率密度函数通常规范化为1,$\boldsymbol{P}_x\{\boldsymbol{\xi}\}$ 趋近于1。概率密度函数和累积分布函数之间关系如下

$$p(x) = \left[\frac{\mathrm{d}}{\mathrm{d}\boldsymbol{\xi}} P_x\{\boldsymbol{\xi}\}\right]_{\xi = x} \tag{2-81}$$

2.4.1.3 联合概率密度函数和边界概率密度函数

两个多维随机变量的联合概率密度函数是联合事件出现的概率 $p(\boldsymbol{x}, z)$。$\boldsymbol{x}$ 的边界概率密度函数定义为

$$p(\boldsymbol{x}) = \int_{-\infty \atop \mathbf{R}^{n_z}}^{\infty} p(\boldsymbol{x}, z)\,\mathrm{d}z \tag{2-82}$$

条件 z 下 $\boldsymbol{x}$ 的条件概率分布是

$$p(\boldsymbol{x} \mid z) = \frac{p(\boldsymbol{x}, z)}{p(z)} \tag{2-83}$$

因此,$\boldsymbol{x}$ 的概率分布函数可以由下式得,即

$$p(\boldsymbol{x}) = \int_{-\infty \atop \mathbf{R}^{n_z}}^{\infty} p(\boldsymbol{x}, z)\,\mathrm{d}z = \int_{-\infty \atop \mathbf{R}^{n_z}}^{\infty} p(\boldsymbol{x} \mid z) p(z)\,\mathrm{d}z \tag{2-84}$$

2.4.1.4 贝叶斯法则

贝叶斯法则服从条件概率

$$p(\boldsymbol{x}, z) = p(\boldsymbol{x} \mid z) p(z) = p(z \mid \boldsymbol{x}) p(\boldsymbol{x}) \tag{2-85}$$

$$p(\boldsymbol{x} \mid z) = \frac{p(z \mid \boldsymbol{x}) p(\boldsymbol{x})}{p(z)} = \frac{p(z \mid \boldsymbol{x}) p(\boldsymbol{x})}{\int_{-\infty \atop \mathbf{R}^{n_x}}^{\infty} p(z \mid \boldsymbol{x}) p(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}} \tag{2-86}$$

由此,有如下比例($\propto$ 表示比例),即

$$p(\boldsymbol{x} \mid z) \propto p(z \mid \boldsymbol{x}) p(\boldsymbol{x}) \tag{2-87}$$

2.4.1.5 概率密度函数的点估计(矩)

假定随机矢量 $\boldsymbol{x} \triangleq [x_1, x_2, \cdots, x_{n_1}]^{\mathrm{T}} \in \mathbf{R}^{n_x}$ 的概率密度函数为 $p(\boldsymbol{x})$。$\boldsymbol{x}$ 的期望值(一阶矩)定义为

$$\hat{\boldsymbol{x}} \triangleq \varepsilon\{\boldsymbol{x}\} = \int_{-\infty \atop \mathbf{R}^{n_x}}^{\infty} \boldsymbol{x}_p(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \tag{2-88}$$

协方差(二阶矩)为

$$\begin{aligned} \boldsymbol{P}^{xx} &\triangleq \varepsilon\{(\boldsymbol{x} - \hat{\boldsymbol{x}})(\boldsymbol{x} - \hat{\boldsymbol{x}})^{\mathrm{T}}\} \\ &= \int_{-\infty \atop \mathbf{R}^{n_x}}^{\infty} (\boldsymbol{x} - \hat{\boldsymbol{x}})(\boldsymbol{x} - \hat{\boldsymbol{x}})^{\mathrm{T}} p(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \end{aligned}$$

$$= \int_{-\infty}^{\infty} \boldsymbol{x}\boldsymbol{x}^{\mathrm{T}} p(\boldsymbol{x})\mathrm{d}\boldsymbol{x} - \hat{\boldsymbol{x}}\hat{\boldsymbol{x}}^{\mathrm{T}} \tag{2-89}$$

对于两个 n_x 维随机矢量 $\boldsymbol{x} \triangleq [x_1, x_2, \cdots, x_{n_x}]^{\mathrm{T}}$ 和 $\boldsymbol{z} \triangleq [z_1, z_2, \cdots, z_{n_x}]^{\mathrm{T}}$，两个变量之间的互协方差定义为

$$\begin{aligned} \boldsymbol{P}^{xz} &\triangleq \varepsilon\{(\boldsymbol{x} - \hat{\boldsymbol{x}})(\boldsymbol{z} - \hat{\boldsymbol{z}})^{\mathrm{T}}\} \\ &= \int_{\substack{-\infty \\ \mathbf{R}^{n_x}}}^{\infty} (\boldsymbol{x} - \hat{\boldsymbol{x}})(\boldsymbol{z} - \hat{\boldsymbol{z}})^{\mathrm{T}} p(\boldsymbol{x}, \boldsymbol{z})\mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{z} \\ &= \int_{-\infty}^{\infty} \boldsymbol{x}\boldsymbol{z}^{\mathrm{T}} p(\boldsymbol{x}, \boldsymbol{z})\mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{z} - \hat{\boldsymbol{x}}\hat{\boldsymbol{z}}^{\mathrm{T}} \end{aligned} \tag{2-90}$$

注意，一些概率密度的矩是不存在的，这是因为这些矩的积分是无限大的。

2.4.2 高斯密度

1733 年，AbrahamdeMoivre[18] 在一篇论文中首次介绍了高斯概率密度函此，自此其在估计理论中的地位就非常重要。因为这种对称分布具有简单的解析形式，对于标量和矢量随机变量都很适用，可以在许多领域中用于噪声过程建模。在线参考文献[19]中总结了高斯分布的属性。下面将讨论高斯密度的特性，这对本书提出的递归估计方法非常重要。首先介绍高斯概率密度函数，然后给出高斯累积分布函数的推导，之后介绍一种仿射变换，将坐标系统的原点变换至高斯密度的均值，旋转坐标轴使其与密度函数的主轴重合。这样做主要是为了简化第 Ⅱ 部分中提出的高斯加权积分方法。最后，给出变换矩方程一组恒等式。

2.4.2.1 高斯概率密度函数

在估计随机变量的矩的过程中，最常用的假设是概率分布符合高斯分布。一维随机变量 x 服从高斯概率密度函数，如果

$$p(x) = \mathcal{N}(x; \hat{x}, \sigma^2) \triangleq \frac{1}{\sqrt{2\pi\sigma}} \exp\left\{-\frac{1}{2\sigma^2}[x - \hat{x}]^2\right\} \tag{2-91}$$

式中：$\hat{x}$ 为一阶矩或者均值；σ^2 为二阶矩或者方差（σ 通常称作标准差）。图 2-5 中给出了关于 x 的一维高斯概率密度函数的示例，其中 $\hat{x} = -5$、$\sigma = 6$。

n_x 维随机变量 $\boldsymbol{x}$ 的高斯概率密度函数定义为

$$\begin{aligned} p(\boldsymbol{x}) = \mathcal{N}(\boldsymbol{x}; \hat{\boldsymbol{x}}, \boldsymbol{P}^{xx}) &\triangleq (2\pi)^{-n_x/2} |\boldsymbol{P}^{xx}|^{-1/2} \cdot \\ &\exp\left\{-\frac{1}{2}[\boldsymbol{x} - \hat{\boldsymbol{x}}]^{\mathrm{T}} (\boldsymbol{P}^{xx})^{-1} [\boldsymbol{x} - \hat{\boldsymbol{x}}]\right\} \end{aligned} \tag{2-92}$$

式中：$\hat{\boldsymbol{x}}$ 和 $\boldsymbol{P}^{xx}$ 分别为矢量均值和协方差矩阵；$|\boldsymbol{A}|$ 为矩阵 $\boldsymbol{A}$ 的行列式。注意 p 始终为一个标量。

假定二维高斯概率密度函数

$$\hat{\boldsymbol{x}} = [\hat{x}, \hat{y}]^{\mathrm{T}} = [-5, 2]^{\mathrm{T}} \tag{2-93}$$

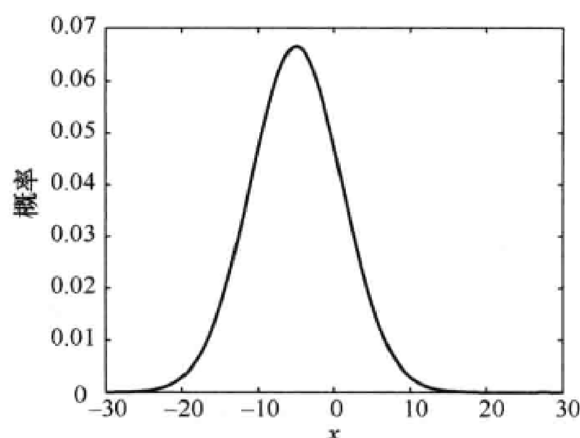


图 2 - 5　一维高斯概率密度函数示例

且 $\sigma_x = 6, \sigma_y = 8, \rho = 0.5$，将 $\boldsymbol{P}^{xx} \rightarrow \boldsymbol{P}^{xy}$，得

$$\boldsymbol{P}^{xy} = \begin{bmatrix} \sigma_x^2 & \rho\sigma_x\sigma_y \\ \rho\sigma_x\sigma_y & \sigma_y^2 \end{bmatrix} = \begin{bmatrix} 6^2 & 24 \\ 24 & 8^2 \end{bmatrix} \tag{2-94}$$

二维互协方差可写成如下形式

$$\boldsymbol{P}^{xy} = \begin{bmatrix} \sigma_x^2 & \sigma_{xy}^2 \\ \sigma_{xy}^2 & \sigma_y^2 \end{bmatrix} \tag{2-95}$$

$$\sigma_{xy}^2 \triangleq \rho\sigma_x\sigma_y$$

图 2 - 6 给出了这种概率密度函数的图形化描述。

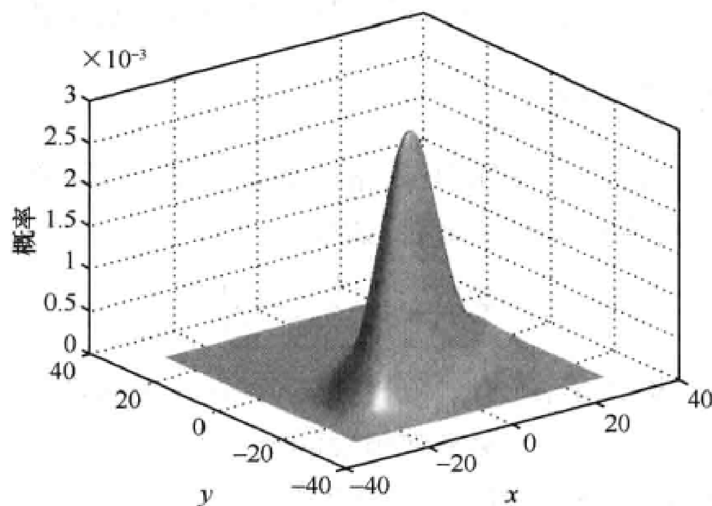


图 2 - 6　二维高斯概率密度函数示例

2.4.2.2　高斯累积分布函数

多维高斯累积分布函数(cdf)的定义为

$$P(\boldsymbol{x}) \triangleq \int_{\substack{-\infty \\ \mathbf{R}^{n_x}}}^{x} p(\boldsymbol{u})\,\mathrm{d}\boldsymbol{u} = \int_{\substack{-\infty \\ \mathbf{R}^{n_x}}}^{x} \mathcal{N}(\boldsymbol{u}; \hat{\boldsymbol{x}}, \boldsymbol{P}^{xx})\,\mathrm{d}\boldsymbol{u}$$

$$= (2\pi)^{-n_x/2} |\boldsymbol{P}^{xx}|^{-1/2} \cdot$$

$$\int_{\substack{-\infty \\ \mathbb{R}^{n_x}}}^{x} \exp\left\{-\frac{1}{2}[\boldsymbol{u} - \hat{\boldsymbol{x}}]^{\mathrm{T}}(\boldsymbol{P}^{xx})^{-1}[\boldsymbol{u} - \hat{\boldsymbol{x}}]\right\}\mathrm{d}\boldsymbol{u} \quad (2-96)$$

令

$$\boldsymbol{t}(\boldsymbol{u}) \triangleq \boldsymbol{D}^{-1}(\boldsymbol{u} - \hat{\boldsymbol{x}}) \quad (2-97)$$

$\boldsymbol{D}$ 可以定义为

$$\boldsymbol{P}^{xx} = \boldsymbol{D}\boldsymbol{D}^{\mathrm{T}}$$

则 $P(\boldsymbol{x})$ 变为

$$P(\boldsymbol{x}) = \prod_{j=1}^{n_x}\left\{\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{t_j(x)} \mathrm{e}^{-t_j^2/2}\mathrm{d}t_j\right\} \quad (2-98)$$

式中:t_j 为 $\boldsymbol{t}$ 的第 j 个元素。括号内的一维方程可重写为

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{t_j(x)} \mathrm{e}^{-t_j^2/2}\mathrm{d}t_j(\boldsymbol{u}) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{0} \mathrm{e}^{-t_j^2/2}\mathrm{d}t_j + \frac{1}{\sqrt{2\pi}}\int_{0}^{t_j} \mathrm{e}^{-t_j^2/2}\mathrm{d}t_j$$

$$= \frac{1}{2} + \frac{1}{2}\left(\frac{2}{\sqrt{\pi}}\int_{0}^{t_j} \mathrm{e}^{-t_j^2/2}\mathrm{d}t_j\right) \quad (2-99)$$

注意 MATLAB® 中的函数定义为

$$\mathrm{erf}[q] = \frac{2}{\sqrt{\pi}}\int_{0}^{q} \mathrm{e}^{-u^2/2}\mathrm{d}u \quad (2-100)$$

此时多维高斯累积分布函数变成

$$P(\boldsymbol{x}) = \prod_{j=1}^{n_x} \frac{1}{2}\{1 + \mathrm{erf}[t_j]\} \quad (2-101)$$

图2-7、图2-8分别给出一维、二维累积分布函数的例子,其中均值和方差的取值和上面一维、二维概率密度函数示例相同。

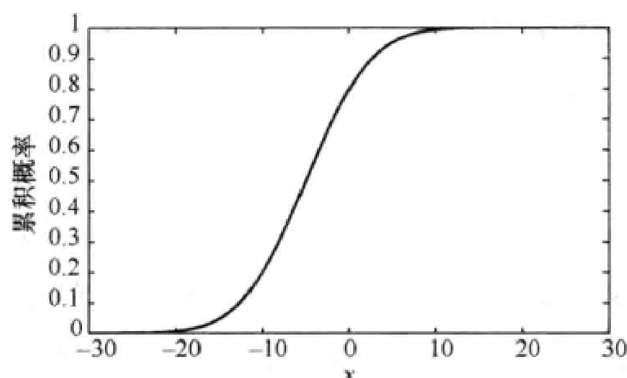


图 2-7　二维高斯概率密度函数示例

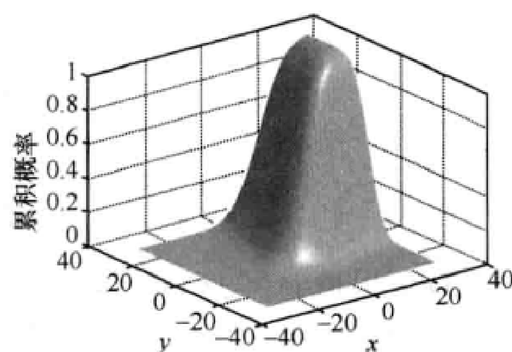


图 2-8　二维高斯累积分布函数的例子

2.4.2.3 一般高斯矩方程的变换

注意矩方程(2 - 88) 和方程(2 - 89) 中,积分的形式为$\int \boldsymbol{x}p(\boldsymbol{x})\mathrm{d}\boldsymbol{x}$。对于一维高斯情况,考虑更通用的矩积分为

$$
\begin{aligned}
\mathcal{E}\{f(x)\} &= \int_{-\infty}^{\infty} f(x)\mathcal{N}(x;\hat{x},\sigma^2)\mathrm{d}x \\
&= \int_{-\infty}^{\infty} f(x)\left(\frac{1}{\sqrt{2\pi}\sigma}\exp\left\{-\frac{1}{2\sigma^2}[x-\hat{x}]^2\right\}\right)\mathrm{d}x
\end{aligned} \tag{2-102}
$$

进行一维仿射变换

$$
x = \hat{x} + \sigma c \tag{2-103}
$$

该变换将坐标原点 $\hat{x}$ 进行了平移并且对标准差进行了比例变换,故 $\mathcal{N}(x;\hat{x},\sigma^2) \to \mathcal{N}(c;0,1)/\sigma$。其中,$c$ 是一个均值为零的单位标准差随机变量。
令

$$
\tilde{f}(c) \triangleq f(\hat{x} + \sigma c) \tag{2-104}
$$

且 $\mathrm{d}x \to \sigma \mathrm{d}c$,一般的矩积分方程(2 - 102) 可变为

$$
\mathcal{E}\{f(x)\} = \int_{-\infty}^{\infty} \tilde{f}(c)\mathcal{N}(c;0,1)\mathrm{d}c \tag{2-105}
$$

类似地,对于多维情况,考虑一般矩积分

$$
\begin{aligned}
\mathcal{E}\{\boldsymbol{f}(\boldsymbol{x})\} &= \int_{\substack{-\infty \\ \mathbb{R}^{n_x}}}^{\infty} \boldsymbol{f}(\boldsymbol{x})\mathcal{N}(\boldsymbol{x};\hat{\boldsymbol{x}},\boldsymbol{P}^{xx})\mathrm{d}\boldsymbol{x} \\
&= \int_{\substack{-\infty \\ \mathbb{R}^{n_x}}}^{\infty} \boldsymbol{f}(\boldsymbol{x})(2\pi)^{-n_x/2}|\boldsymbol{P}^{xx}|^{-1/2} \cdot \\
&\quad \exp\left\{-\frac{1}{2}[\boldsymbol{x}-\hat{\boldsymbol{x}}]^{\mathrm{T}}(\boldsymbol{P}^{xx})^{-1}[\boldsymbol{x}-\hat{\boldsymbol{x}}]\right\}\mathrm{d}\boldsymbol{x}
\end{aligned} \tag{2-106}
$$

现在进行多维仿射变换

$$
\boldsymbol{x} = \hat{\boldsymbol{x}} + \boldsymbol{D}\boldsymbol{c} \tag{2-107}
$$

$$
\boldsymbol{P}^{xx} = \boldsymbol{D}\boldsymbol{D}^{\mathrm{T}} \tag{2-108}
$$

得恒等式

$$
\boldsymbol{D} = [\boldsymbol{P}^{xx}]^{1/2} \tag{2-109}
$$

这里,$\boldsymbol{D}$ 可以取 $\boldsymbol{P}^{xx}$ 的 Cholesky 因子或 $\boldsymbol{P}^{xx}$ 主(正定)矩阵的根。实际上,任何矩阵 $\boldsymbol{D}$ 只要满足式(2 - 108) 中的定义就足够了。此外,$\boldsymbol{c}$ 可以看成是具有单位对角协方差矩阵的零均值多维随机变量,即 $\boldsymbol{c} \sim \mathcal{N}(\boldsymbol{c};\boldsymbol{0},\boldsymbol{I})$。

式(2 - 107) 中的仿射变换有一个几何解释。多维高斯分布的等概率轮廓的是以均值为中心的椭球体。主轴的方向由协方差矩阵的特征矢量决定,特征值等

于主轴相对长度的平方。式(2 - 107) 所示变换将坐标原点平移至 $\hat{\boldsymbol{x}}$ 坐标,并将坐标轴旋转至与密度函数的主轴重合,同时将每个轴比例变换至标准差为 1,如图 2 - 9 所示。注意,转换后的密度的轮廓是圆形而不是椭圆形。

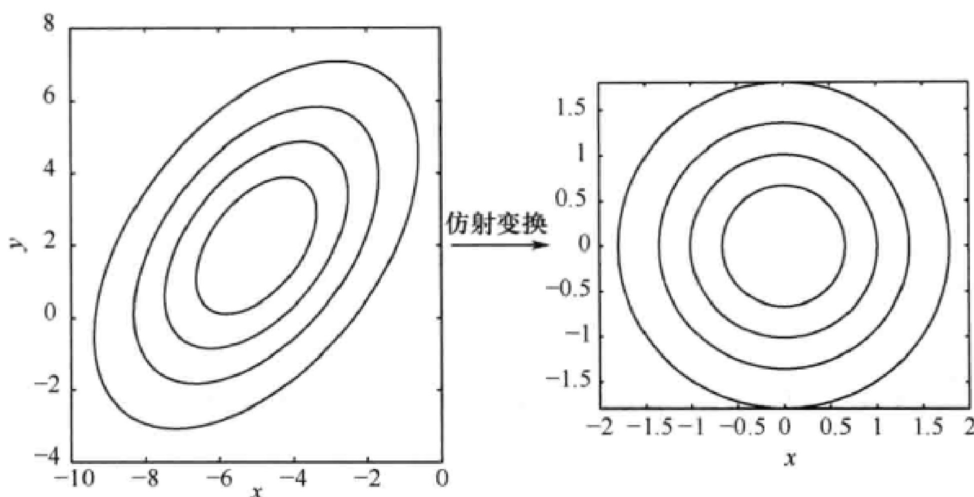


图 2 - 9 对高斯概率密度函数进行仿射变换的效果

考虑到

$$\int_{-\infty \atop \mathbf{R}^{n_x}}^{\infty} \boldsymbol{f}(\boldsymbol{x}) \mathcal{N}(\boldsymbol{x};\hat{\boldsymbol{x}},\boldsymbol{P}^{xx}) \mathrm{d}\boldsymbol{x} = \int_{-\infty \atop \mathbf{R}^{n_x}}^{\infty} \boldsymbol{f}(\boldsymbol{x}(\boldsymbol{c})) \mathcal{N}(\boldsymbol{x};\hat{\boldsymbol{x}},\boldsymbol{P}^{xx}) \left| \frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}\boldsymbol{c}} \right| \mathrm{d}\boldsymbol{c} \qquad (2-110)$$

将式(2 - 110) 和式(2 - 107) 代入式(2 - 106),得

$$\mathcal{E}\{\boldsymbol{f}(\boldsymbol{x})\} = \int_{-\infty \atop \mathbf{R}^{n_x}}^{\infty} \tilde{\boldsymbol{f}}(\boldsymbol{c}) \mathcal{N}(\boldsymbol{c};\boldsymbol{0},\boldsymbol{I}) \mathrm{d}\boldsymbol{c} \qquad (2-111)$$

$$\tilde{\boldsymbol{f}}(\boldsymbol{c}) \triangleq \boldsymbol{f}(\hat{\boldsymbol{x}} + \boldsymbol{D}\boldsymbol{c}) \qquad (2-112)$$

这种仿射变换的应用是一个非常重要的概念,可将其用于简化高斯加权积分的数值估计方法,建立贝叶斯估计的矩方程组。使用这种仿射变换,结合第 2.3 节中讨论的非线性函数的多项式展开,简化了这些积分的数值估计过程。

2.4.2.4 高斯矩方程

本节将介绍(并不推导) 一些适用于高斯矩方程的恒等式。

对于一个随机变量 $\boldsymbol{x} \in \boldsymbol{R}^{n_x}$ 和一个对称矩阵 $\boldsymbol{\Lambda} \in \boldsymbol{R}^{n_x \times n_x}$,标量化数量 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\Lambda}\boldsymbol{x}$ 被称为 $\boldsymbol{x}$ 的二次型。可以表示为[20,21]

$$\mathcal{E}\{\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\Lambda}\boldsymbol{x}\} = \mathrm{tr}\{\boldsymbol{\Lambda}\boldsymbol{P}\} + \hat{\boldsymbol{x}}^{\mathrm{T}}\boldsymbol{\Lambda}\hat{\boldsymbol{x}} \qquad (2-113)$$

式中:$\hat{\boldsymbol{x}}$ 和 $\boldsymbol{P}$ 分别是 $\boldsymbol{x}$ 的期望值和协方差矩阵;{·} 表示括号内矩阵的迹。这一结果仅需要 $\hat{\boldsymbol{x}}$ 和 $\boldsymbol{P}$ 存在,而与概率密度函数的形式无关。用 $\boldsymbol{x}$ 的四次形式表达期望值如下[20]

$$\mathcal{E}\{\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\Lambda}_1\boldsymbol{x}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\Lambda}_2\boldsymbol{x}\} = \mathrm{tr}\{\boldsymbol{\Lambda}_1\boldsymbol{P}\}\mathrm{tr}\{\boldsymbol{\Lambda}_2\boldsymbol{P}\} + 2\mathrm{tr}\{\boldsymbol{\Lambda}_1\boldsymbol{P}\boldsymbol{\Lambda}_2\boldsymbol{P}\} \qquad (2-114)$$

在一维情况下,高斯矩方程包含积分项 $\int_{-\infty}^{\infty} c^n \mathcal{N}(c;0,1)\mathrm{d}c$。由于 $\hat{\boldsymbol{c}} = 0$,并且 $\mathcal{N}(c;$

0,1) 关于 0 对称，得的矩方程 c^n 如下

$$\mathcal{E}\{c^n\} = \int_{-\infty}^{\infty} c^n \mathcal{N}(c;0,1)\mathrm{d}c = \begin{cases} 1, & n=0 \\ 1\times 3\times 5\cdots(n-1), & n \text{ 为奇数} \\ 0, & n \text{ 为偶数} \end{cases} \tag{2-115}$$

根据式(2 - 115) 的一维矩方程，得 $\boldsymbol{c} = [c_1, c_2, \cdots, c_{n_x}]^{\mathrm{T}}$ 中分量的四阶矩方程如下

$$\int \mathcal{N}(\boldsymbol{c};\boldsymbol{0},\boldsymbol{I})\mathrm{d}\boldsymbol{c} = \int_{-\infty}^{\infty}\mathcal{N}(c_1;0,1)\mathrm{d}c_1\cdots\int_{-\infty}^{\infty}\mathcal{N}(c_{n_x};0,1)\mathrm{d}c_{n_x} = 1 \tag{2-116}$$

$$\int c_i\mathcal{N}(\boldsymbol{c};\boldsymbol{0},\boldsymbol{I})\mathrm{d}\boldsymbol{c} = \int c_i\mathcal{N}(c_i;0,1)\mathrm{d}c_i\int\mathcal{N}(\boldsymbol{c}^*;\boldsymbol{0},\boldsymbol{I})\mathrm{d}\boldsymbol{c}^* = 0 \tag{2-117}$$

$$\int c_ic_j\mathcal{N}(\boldsymbol{c};\boldsymbol{0},\boldsymbol{I})\mathrm{d}\boldsymbol{c} = \delta_{ij} \tag{2-118}$$

$$\begin{aligned}\int c_ic_jc_k\mathcal{N}(\boldsymbol{c};\boldsymbol{0},\boldsymbol{I})\mathrm{d}\boldsymbol{c} &= \delta_{ij}\int c_k\mathcal{N}(c_k;0,1)\mathrm{d}c_k + \\ &\quad\ \delta_{ik}\int c_j\mathcal{N}(c_j;0,1)\mathrm{d}c_j + \\ &\quad\ \delta_{jk}\int c_i\mathcal{N}(c_i;0,1)\mathrm{d}c_i \\ &= 0\end{aligned} \tag{2-119}$$

$$\int c_i^2c_j^2\mathcal{N}(\boldsymbol{c};\boldsymbol{0},\boldsymbol{I})\mathrm{d}\boldsymbol{c} = \begin{cases}1, & i\neq j\\ 3, & i=j\end{cases} \tag{2-120}$$

$$\int c_ic_jc_kc_l\mathcal{N}(\boldsymbol{c};\ \boldsymbol{0},\boldsymbol{I})\mathrm{d}\boldsymbol{c} = \delta_{ij,kl} + \delta_{ik,jl} + \delta_{il,jk} + 3\delta_{ijkl} \tag{2-121}$$

$$\boldsymbol{c}^* \triangleq [c_1,\cdots,c_{i-1},c_{i+1},\cdots,c_{n_x}]^{\mathrm{T}}$$

为计算方便，去掉所有积分的隐式积分极限，得

$$\delta_{ij} \triangleq \begin{cases}0, & i\neq j\\ 1, & i=j\end{cases} \tag{2-122}$$

$$\delta_{ijkl} \triangleq \begin{cases}1, & i=j=k=l\\ 0, & \text{其他}\end{cases} \tag{2-123}$$

$$\delta_{ij,kl} \triangleq \begin{cases}1, & i=j,k=l;\ i,j\neq k,l\\ 0, & \text{其他}\end{cases} \tag{2-124}$$

$\delta_{ik}\delta_{jl}$ 和 $\delta_{il}\delta_{jk}$ 的定义与上面类似。

由式(2 - 113) 和式(2 - 114)，得 $\boldsymbol{c}$ 的二次矩方程如下

$$\int[\boldsymbol{c}^{\mathrm{T}}\boldsymbol{M}\boldsymbol{c}]^i\mathcal{N}(\boldsymbol{c};\ \boldsymbol{0},\boldsymbol{I})\mathrm{d}\boldsymbol{c} = \begin{cases}1, & i=0\\ \mathrm{tr}\{\boldsymbol{M}\}, & i=1\\ [\mathrm{tr}\{\boldsymbol{M}\}]^2 + 2\mathrm{tr}\{\boldsymbol{M}\boldsymbol{M}\}, & i=2\end{cases} \tag{2-125}$$

式中:$\boldsymbol{M} \in \mathbf{R}^{n_x \times n_x}$ 为一已知对称方阵。

根据式(2 - 116) 至式(2 - 121),很容易证明

$$\int \boldsymbol{c}\boldsymbol{c}^{\mathrm{T}} \mathcal{N}(\boldsymbol{c};\ \boldsymbol{0},\boldsymbol{I})\,\mathrm{d}\boldsymbol{c} = \boldsymbol{I} \tag{2-126}$$

2.4.2.5 使用 MATLAB® 从多维高斯分布中生成采样

一般多维高斯分布 $\mathcal{N}(\boldsymbol{x};\hat{\boldsymbol{x}},\boldsymbol{P}^{xx})$ 中的采样可以在 MATLAB 中采用如下方式生成:

(1) 生成采样点$\{\boldsymbol{c}^{(i)} \sim \mathcal{N}(\boldsymbol{c};\boldsymbol{0},\boldsymbol{I}), i = 1,\cdots,n_x\}$。这里,符号 ~ 代表从密度中选取的第 i 个采样 $\mathcal{N}(\boldsymbol{c};\boldsymbol{0},\boldsymbol{I})$。在 MATLAB 中,使用命令 $c = randn(N,M)$ 来完成采样,其中 N 为 $\boldsymbol{c}$ 的维度,M 是所需的采样个数。

(2) 使用式(2 - 107) 将对 $\boldsymbol{c}$ 的采样变换为对 $\boldsymbol{x}$ 的采样,得

$$\boldsymbol{x}^{(i)} = \hat{\boldsymbol{x}} + \boldsymbol{D}\boldsymbol{c}^{(i)} \tag{2-127}$$

式中:$\boldsymbol{D}$ 为 $\boldsymbol{P}^{xx} = \boldsymbol{D}\boldsymbol{D}^{\mathrm{T}}$。

参考文献

[1] Brookes M. The Matrix Reference Manual. [May 16,2011.]*Available online at* http://www. ee. ic. ac. uk/hp/staff/dmb/matrix/intro. html.

[2] Lerner UN. Hybrid Bayesian Networks for Reasoning About Complex Systems,Dissertation. Stanford, CA: Stanford University,2002.

[3] McNamee J,Stenger F. Construction of fully symmetric numerical integration formulas. Numerische Math,1967,10:327 - 344.

[4] Arfken G. *Mathematical Methods for Physicists.* 3rd ed. Academic Press,1985.

[5] Ralston A,Rabinowitz P. *A First Course in Numerical Analysis.* 2nd ed. Dover Press,2001.

[6] Williamson RE,Cromwell RH,Trotter HF. *Calculus of Vector Functions.* 3rd ed. Prentice - Hall,1972.

[7] Julier S,Uhlmann,J. A General Method for Approximating Nonlinear Transformations of Probability Distributions. Department of Engineering Science,University of Oxford,1996.

[8] Norgaard M,Poulsen NK,Ravn O. New developments in state estimation for nonlinear systems. *Automatica*, 2000,36:1627 - 1638.

[9] Burden,RL,Faires,JD. *Numerical Analysis.* Belmont: Brooks/Cole,2000.

[10] Silvester P,Ferrari R. *Finite Elements for Electrical Engineers.* 3rd ed. Cambridge University Press,1996.

[11] Waldron S. The error in linear interpolation at the vertices of a simplex. SIAM *J. Numer. Analy*,1998, 35(3):1191 - 1200.

[12] Gasca M,Sauer T. Polynomial interpolation in several variables. *Adv. Comput. Math*,2000,12(4): 377 - 410.

[13] Stampfle M. Optimal estimates for the linear interpolation error on simplices. *J. Approx. Theory*, 2000,

103:78 - 90.

[14] Lu Y, Wang RH. A multidimensional generalization of Simpson - type formulas over Nsimplex. *J. Inform. Comput. Sci*, 2004, 1(2):331 - 335.

[15] Hemingway P. n - Simplex Interpolation. HPL - 2002 - 320, Client and Media Systems Laboratory, HP laboratories, Bristol, November 2002.

[16] Warburton T. An explicit construction for interpolation nodes on the simplex. *J. Eng. Math*, 2006, 56(3):247 - 262.

[17] *Papoulis A. Probability, Random Variables, and Stochastic Processes. McGraw - Hill*, 1965.

[18] deMoivre A. Approximatio ad Summam Terminorum Binomii $(a + b)n$ in Seriem expansi. November 12, 1733.

[19] [online] http://en. wikipedia. org/wiki/Normal distribution.

[20] Bar Shalom Y, Li XR, Kirubarajan T. *Estimation with Application to Tracking and Navigation: Theory, Algorithms and Software.* Wiley, 2001.

[21] Koch KR. *Introduction to Bayesian Statistics.* New York: Springer, 2007.

第3章 贝叶斯估计的基本概念

估计过程通过实验得一组观测样本，通常是某种形式的数据。观测数据可以是对某个特定目标的方位和(或)范围的连续时间采样，或者是一个交易日内股票价格的连续采样。根据观测数据，可以对一些无法直接观测到的特征参数进行估计。例如，通过方位观测，可以用时间相关的函数估计出目标位置和速度；而对于股票价格数据，则可以估计出股票的波动。

本书集中讨论了多种估计方法，包括随时间动态变化的变量的模型估计、可观测数据和不可观测变量之间关系的模型估计。特别指出，本书的讨论将局限于离散采样数据的递推估计方法。因此，书中假定待估计参数服从已知的递推动态过程，而且观测数据和估计参数值之间的解析关系已知。

此外，贝叶斯估计假设待估计参数和观测数据都是随机的。观测数据和待估计参数之间的解析关系(一种变换)为估计提供了一个统一的框架，其中递归推理以当前和以前所有观测作为已知条件，当前状态矢量的密度函数来表示。

3.1 贝叶斯估计

贝叶斯估计的目标是，给定一个观测矢量$\boldsymbol{z}$，估计参数矢量$\boldsymbol{x}$的连续值。一般将$\boldsymbol{x}$和$\boldsymbol{z}$作为随机矢量来处理。对于参数矢量，控制参数动态变化的方程本身就有随机假设，其中添加未建模影响因子作为随机噪声。假设随机测量噪声存在，可以证明观测矢量的随机性质。假设随机矢量$\boldsymbol{x}$有已知的先验密度函数$p(\boldsymbol{x})$，先验分布包括在得任何可观测数据之前的已知和未知的所有参数矢量。如果参数$\boldsymbol{x}$的真值已知，那么$\boldsymbol{z}$的概率密度可由条件密度或似然函数$p(\boldsymbol{z}|\boldsymbol{x})$表示，$\boldsymbol{z}$的完整统计特性已知。

曾经进行过这样一项试验，随机变量$\boldsymbol{z}$已知，可通过贝叶斯法则得$\boldsymbol{x}$的后验

条件密度为

$$p(\boldsymbol{x} \mid z) = \frac{p(z \mid \boldsymbol{x})p(\boldsymbol{x})}{p(z)} \tag{3-1}$$

因此,在贝叶斯框架下,后验密度 $p(\boldsymbol{x} \mid z)$ 中包括考虑试验观测结果之后,需要了解的 $\boldsymbol{x}$ 的所有信息。由于试验结果 z 已知,式(3-1)中的分母就是一个归一化标量常数,可以得

$$p(z) = \int_{\mathbf{R}^{n_x}} p(z \mid \boldsymbol{x})p(\boldsymbol{x})\mathrm{d}\boldsymbol{x} \tag{3-2}$$

对于完全贝叶斯估计问题的似然密度和后验密度,或者它们的联合密度,定义估计的统计模型,其中参数的联合密度和观测矢量定义为

$$p(\boldsymbol{x},z) = p(\boldsymbol{x} \mid z)p(z) \tag{3-3}$$

基于后验分布估计问题的解决方案如式(3-1),因此后验分布可以按需要生成 $\boldsymbol{x}$ 的点估计(如果存在点估计)。注意,后验密度被认为是估计问题最一般的解决方案,在很多情况下可以用密度函数来描述 $\boldsymbol{x}$。

3.2 点估计式

后验密度的评价是一个非常困难的问题,是多维推理问题的复杂解。点估计 $\hat{\boldsymbol{x}}$ 是在给定观测值情况下,对参数值进行有根据的推测。基于所有得观测值,生成点估计 $\hat{\boldsymbol{x}}$ 的解析方法称为 $\boldsymbol{x}$ 的估计器。使用的估计器不同,得的 $\hat{\boldsymbol{x}}$ 的实际值不同。即便使用同样的估计器,每次试验得的值也不同。因此,点估计本身被当做一个随机变量。

为找到合适的估计器,定义一个代价函数 $L(\hat{\boldsymbol{x}},\boldsymbol{x})$,对于错误的估计 $\hat{\boldsymbol{x}} \neq \boldsymbol{x}$ 进行惩罚。通常,代价函数的惩罚是根据误差的大小 $\boldsymbol{x}-\hat{\boldsymbol{x}}$ 确定的。在不失一般性的情况下,假设代价函数为正,$L(\boldsymbol{x},\boldsymbol{x}) = 0$ 是其唯一最小值。对于点估计 $\hat{\boldsymbol{x}}$,贝叶斯风险 R 为代价函数的期望值

$$R \triangleq \mathcal{E}\{L(\hat{\boldsymbol{x}},\boldsymbol{x})\} \tag{3-4}$$

根据 $\boldsymbol{x}$ 和 z 求得期望值。注意,使用观测值 z 来生成 $\hat{\boldsymbol{x}}$,可写为 $\hat{\boldsymbol{x}}(z)$。$\hat{\boldsymbol{x}}(z)$ 的最优解是使贝叶斯风险最小的值

$$\hat{\boldsymbol{x}}(z) = \arg\min_{\boldsymbol{x}^*(z)} \int_{\mathbf{R}^{n_z+n_x}} L(\boldsymbol{x}^*(z),\boldsymbol{x})p(\boldsymbol{x},z)\mathrm{d}\boldsymbol{x}\mathrm{d}z \tag{3-5}$$

这里的最小值是通过所有的 $\boldsymbol{x}^*(z)$ 值求得,z 和 $\boldsymbol{x}$ 的维数分别是 n_z 和 n_x。

将式(3-3)代入式(3-4),把积分项分成两部分,即

$$\hat{\boldsymbol{x}}(z) = \arg\min_{\boldsymbol{x}^*(z)} \int_{\mathbf{R}^{n_z}} \left[\int_{\mathbf{R}^{n_x}} L(\boldsymbol{x}^*(z),\boldsymbol{x})p(\boldsymbol{x} \mid z)\mathrm{d}\boldsymbol{x} \right] p(z)\mathrm{d}z \tag{3-6}$$

由于 $L(\boldsymbol{x}^*,\boldsymbol{x})$ 和 $p(\boldsymbol{x} \mid z)$ 都是正数,给定任何 z 值,内积分均为正。对于每一个 z,

因为 $p(z)$ 也是正数，使风险最小的 $\boldsymbol{x}^*(z)$ 值也使内积分最小，即

$$\hat{\boldsymbol{x}}(z) = \arg\min_{\boldsymbol{x}^*(z)} \int_{\mathbf{R}^{n_x}} L(\boldsymbol{x}^*(z), \boldsymbol{x}) p(\boldsymbol{x} \mid z) \mathrm{d}\boldsymbol{x} \tag{3-7}$$

式(3－7)定义的贝叶斯估计问题中，代价函数不同，产生的估计器也各不相同。

还有另外一种方法——最大代价最小化（最小最大估计器）可选作估计器，例如

$$\hat{\boldsymbol{x}}(z) = \arg\min_{\boldsymbol{x}^*(z)} \max\{L(\boldsymbol{x}^*(z), \boldsymbol{x})\} \tag{3-8}$$

对于大多数估计问题，假定代价是基于下面的估计误差

$$\boldsymbol{\epsilon}^x \triangleq \boldsymbol{x} - \hat{\boldsymbol{x}} \tag{3-9}$$

则代价函数 $L(\boldsymbol{\epsilon}^x)$ 是一个单变量函数。

（1）一种有代表性的代价函数是加权平方误差代价函数

$$L(\boldsymbol{\epsilon}^x) = \boldsymbol{\epsilon}^{xT} \boldsymbol{M} \boldsymbol{\epsilon}^x \tag{3-10}$$

它强调了较大误差的影响。这里，$\boldsymbol{M}$ 是已知的平方加权矩阵，不依赖于 $\boldsymbol{x}$。

（2）使用绝对误差值代价函数

$$L(\boldsymbol{\epsilon}^x) = |\boldsymbol{\epsilon}^x| \tag{3-11}$$

（3）代价函数在一些区间取值相等而在其余区间取值为零。例如，设置 $L(\boldsymbol{\epsilon}^x)$ 如下

$$L(\boldsymbol{\epsilon}^x) = \begin{cases} 0, & |\boldsymbol{\epsilon}^x| \leqslant \Delta/2 \\ 1, & |\boldsymbol{\epsilon}^x| > \Delta/2 \end{cases} \tag{3-12}$$

具体代价函数的选择取决于估算方法要实现的结果。通常，很难为一个主观的量指定一个解析测度。如上所述，选择代价函数的目标是要找到一个估计值，使代价的期望值最小。

考虑最小化加权平方误差代价函数的期望值，通常称为均方误差准则。在式(3－9)中，用 $\boldsymbol{x}^*$ 替换 $\hat{x}$，将结果代入式(3－10)，则式(3－7)可变为

$$\hat{\boldsymbol{x}}(z) = \arg\min_{\boldsymbol{x}^*(z)} \int_{\mathbf{R}^{n_x}} (\boldsymbol{x} - \boldsymbol{x}^*)^{\mathrm{T}} \boldsymbol{M} (\boldsymbol{x} - \boldsymbol{x}^*) p(\boldsymbol{x} \mid z) \mathrm{d}\boldsymbol{x} \tag{3-13}$$

为得参数的最小值，关于 $\boldsymbol{x}^*$ 求导，将结果设成0，求解 $\boldsymbol{x}^*$，则可以得

$$\begin{aligned} &\frac{\mathrm{d}}{\mathrm{d}\boldsymbol{x}^*} \int_{\mathbf{R}^{n_x}}^{T} (\boldsymbol{x} - \boldsymbol{x}^*)^{\mathrm{T}} \boldsymbol{M} (\boldsymbol{x} - \boldsymbol{x}^*) p(\boldsymbol{x} \mid z) \mathrm{d}\boldsymbol{x} \\ &= -2\boldsymbol{M} \int_{\mathbf{R}^{n_x}} \boldsymbol{x} p(\boldsymbol{x} \mid z) \mathrm{d}\boldsymbol{x} + 2\boldsymbol{M}\boldsymbol{x}^* \int_{\mathbf{R}^{n_x}} p(\boldsymbol{x} \mid z) \mathrm{d}\boldsymbol{x} \end{aligned} \tag{3-14}$$

将它置为0，并将第二个积分项置为1，得

$$\hat{\boldsymbol{x}}_{\mathrm{ms}} = \int_{\mathbf{R}^{n_x}} \boldsymbol{x} p(\boldsymbol{x} \mid z) \mathrm{d}\boldsymbol{x} \triangleq \mathcal{E}\{\boldsymbol{x} \mid z\} \tag{3-15}$$

因此，均方误差代价函数最小化的结果是与后验密度（或条件均值）相关的平均值。

由于均方估计器得的估计是随机变量，要量化这些估计的不确定性，一种简便的不确定性测度由条件协方差矩阵表示，即

$$\begin{aligned}\boldsymbol{P}^{xx} &= \mathcal{E}\{(\boldsymbol{x}-\hat{\boldsymbol{x}})(\boldsymbol{x}-\hat{\boldsymbol{x}})^{\mathrm{T}} \mid \boldsymbol{z}\} \\ &= \int_{\mathbf{R}^{n_x}} (\boldsymbol{x}-\hat{\boldsymbol{x}})(\boldsymbol{x}-\hat{\boldsymbol{x}})^{\mathrm{T}} p(\boldsymbol{x} \mid \boldsymbol{z})\,\mathrm{d}\boldsymbol{x}\end{aligned} \tag{3-16}$$

必须注意的是，对于一些概率密度函数，不存在一个或者更多的点估计。柯西分布是一种没有均值、协方差或高阶矩的分布。然而，它的众数和中位数都有明确定义，协方差可以用离差度量来代替。

本书致力于利用平均平方代价函数的估计方法。绝对误差和统一的代价函数可以导出最大后验(MAP)估计器最大似然估计则是，MAP估计的一个特例。这些估计器的完整推导和讨论可参阅文献[1]中第56-58页。

3.3 基于概率密度函数的递推贝叶斯滤波器

离散动态过程是系统的当前状态依赖于之前一个或多个状态的过程。在连续过程中，当前状态对前一个状态的依赖可以用微分方程来表示。对于离散时间观测，只能在离散时间上进行以观测为条件的估计，微分方程可以用与之相等的有限差分代替，并将观测时间为 t_n 时的状态和在 t_n 之前的观测状态联系起来。

一阶马尔可夫过程中当前状态只取决于上一个状态。因此，可将离散随机马尔可夫动态过程表示为

$$\boldsymbol{x}_n = \boldsymbol{f}_{n-1}(\boldsymbol{x}_{n-1}, \boldsymbol{u}_n, \boldsymbol{\eta}_{n-1}) = \boldsymbol{f}_{n-1}(\boldsymbol{x}_{n-1}) + \boldsymbol{u}_n + \boldsymbol{v}_{n-1} \tag{3-17}$$

式中：$\boldsymbol{x}_n$ 为 t_n 时刻的系统状态(状态$\boldsymbol{f}_{n-1}$)，为确定性转移函数(矩阵)，实现从 t_{n-1} 时刻到 t_n 时刻状态 $\boldsymbol{x}$ 的改变；$\boldsymbol{u}_n$ 为已知的(通常是确定性的)控制，构成驱动动态系统的外部输入。

尽管白噪声 $\boldsymbol{\eta}$(不一定是高斯)可以在输入时开始出现，并通过转移函数变换，但通常假定噪声是加性的，用来表示真实转移函数中未建模的部分。注意上面给出的马尔可夫过程只是真实转移函数的模拟。例如，如果将一艘移动的船舶的运动建模为匀速运动，噪声项代表波浪作用导致直线路径微小偏离的加速度。

关注的具体问题是对通常不可观测的状态矢量 $\boldsymbol{x}_n$ 的估计基于所有试验观测矢量 $z_{1:n} \triangleq \{z_1, z_2, \cdots, z_n\}$ 的集合。假定 t_n 时刻的观测矢量和 t_n 时刻的状态矢量之间的解析关系已知，表示为

$$z_n = h_n(\boldsymbol{x}_n, \boldsymbol{\mu}_n) = h_n(\boldsymbol{x}_n) + \boldsymbol{w}_n \tag{3-18}$$

式中：z_n 为观测矢量；$\boldsymbol{h}_n$ 为确定性观测函数，给出状态矢量和观测矢量的关系。白噪声 $\boldsymbol{\mu}_n$ 不一定是高斯噪声)可以通过转移函数变换，但通常假定观测噪声是加性的。

对于多数感兴趣的问题，$\boldsymbol{f}_{n-1}$ 和 $\boldsymbol{h}_n$ 是缓变的，随时间的变化非常缓慢。因此，

可以认为它们是与时间无关的，即，$\boldsymbol{f}_{n-1} \to \boldsymbol{f}$ 和 $\boldsymbol{h}_{n-1} \to \boldsymbol{h}$。

式(3－17)和式(3－18)所采用的特殊解析形式代表了需要状态矢量估计信息的系统完整建模。在贝叶斯意义上，对 $\boldsymbol{x}_n$ 基于完整观测集合的估计 $\boldsymbol{z}_{1:n}$ 可以通过把问题转化成条件后验密度 $p(\boldsymbol{x}_n \mid \boldsymbol{z}_{1:n})$ 的估计问题来加以理解。本书目的在于提出在这种条件分布和后验分布下与 $\boldsymbol{x}_n$ 相关时刻递推贝叶斯估计方法。

对于后验分布 $p(\boldsymbol{x}_n \mid \boldsymbol{z}_{1:n})$，将式(3－1)的贝叶斯法则改写为

$$p(\boldsymbol{x}_n \mid \boldsymbol{z}_{1:n}) = \frac{p(\boldsymbol{z}_{1:n} \mid \boldsymbol{x}_n)p(\boldsymbol{x}_n)}{p(\boldsymbol{z}_{1:n})} \tag{3-19}$$

后验概率密度函数 $p(\boldsymbol{x}_n \mid \boldsymbol{z}_{1:n})$，是 $\boldsymbol{x}_n$ 的概率密度函数，条件是所有的观测值，包括当前的观测。因为集合$\{\boldsymbol{z}_{1:n}\}$可以写为$(\boldsymbol{z}_n, \boldsymbol{z}_{1:n-1}\}$，式(3－19)可以扩展和分解为以下形式，即

$$p(\boldsymbol{x}_n \mid \boldsymbol{z}_{1:n}) = \frac{p(\boldsymbol{z}_n, \boldsymbol{z}_{1:n-1} \mid \boldsymbol{x}_n)p(\boldsymbol{x}_n)}{p(\boldsymbol{z}_n, \boldsymbol{z}_{1:n-1})} \tag{3-20}$$

将式(3－3)代入分子和分母，变换为

$$p(\boldsymbol{x}_n \mid \boldsymbol{z}_{1:n}) = \frac{p(\boldsymbol{z}_n \mid \boldsymbol{z}_{1:n-1}, \boldsymbol{x}_n)p(\boldsymbol{z}_{1:n-1} \mid \boldsymbol{x}_n)p(\boldsymbol{x}_n)}{p(\boldsymbol{z}_n \mid \boldsymbol{z}_{1:n-1})p(\boldsymbol{z}_{1:n-1})} \tag{3-21}$$

将贝叶斯法则(3－1)应用到 $p(\boldsymbol{z}_{1:n-1} \mid \boldsymbol{x}_n)$，化简结果方程，得

$$p(\boldsymbol{x}_n \mid \boldsymbol{z}_{1:n}) = \frac{p(\boldsymbol{z}_n \mid \boldsymbol{z}_{1:n-1}, \boldsymbol{x}_n)p(\boldsymbol{x}_n \mid \boldsymbol{z}_{1:n-1})p(\boldsymbol{z}_{1:n-1})p(\boldsymbol{x}_n)}{p(\boldsymbol{z}_n \mid \boldsymbol{z}_{1:n-1})p(\boldsymbol{z}_{1:n-1})p(\boldsymbol{x}_n)}$$

$$= \frac{p(\boldsymbol{z}_n \mid \boldsymbol{z}_{1:n-1}, \boldsymbol{x}_n)p(\boldsymbol{x}_n \mid \boldsymbol{z}_{1:n-1})}{p(\boldsymbol{z}_n \mid \boldsymbol{z}_{1:n-1})} \tag{3-22}$$

$$= \frac{p(\boldsymbol{z}_n \mid \boldsymbol{x}_n)p(\boldsymbol{x}_n \mid \boldsymbol{z}_{1:n-1})}{p(\boldsymbol{z}_n \mid \boldsymbol{z}_{1:n-1})} \tag{3-23}$$

其中 $p(\boldsymbol{z}_n \mid \boldsymbol{z}_{1:n-1}\boldsymbol{x}_n) \to p(\boldsymbol{z}_n \mid \boldsymbol{x}_n)$，由式(3－18)可知，$t_n$ 时刻的观测值不依赖于 $t_{1:n-1}$ 时刻的观测值。

最后创建一个完全递推形式的条件概率密度函数方程。Chapman－Kolmogorov 方程提供了先验概率密度 $p(\boldsymbol{x}_n \mid \boldsymbol{z}_{1:n-1})$ 与后验密度之间的联系，即

$$p(\boldsymbol{x}_n \mid \boldsymbol{z}_{1:n-1}) = \int p(\boldsymbol{x}_n \mid \boldsymbol{x}_{n-1}, \boldsymbol{z}_{1:n-1})p(\boldsymbol{x}_{n-1} \mid \boldsymbol{z}_{1:n-1})\mathrm{d}\boldsymbol{x}_{n-1}$$

$$= \int p(\boldsymbol{x}_n \mid \boldsymbol{x}_{n-1})p(\boldsymbol{x}_{n-1} \mid \boldsymbol{z}_{1:n-1})\mathrm{d}\boldsymbol{x}_{n-1} \tag{3-24}$$

由式(3－24)和式(3－23)可知，建立起前一个后验概率 $p(\boldsymbol{x}_{n-1} \mid \boldsymbol{z}_{1:n-1})$ 和当前后验概率 $p(\boldsymbol{x}_n \mid \boldsymbol{z}_{1:n})$ 之间的递推关系，需要由 $p(\boldsymbol{x}_n \mid \boldsymbol{x}_{1:n-1})$ 给定预测密度和似然函数 $p(\boldsymbol{z}_n \mid \boldsymbol{x}_n)$。

在式(3－24)中，$p(\boldsymbol{x}_n \mid \boldsymbol{x}_{1:n-1})$ 替换了 $p(\boldsymbol{x}_n \mid \boldsymbol{x}_{n-1}, \boldsymbol{z}_{1:n-1})$，因为预测密度是由式(3－17)动态方程定义的，不依赖于 $\boldsymbol{z}_{1:n-1}$。式(3－17)表明，$\boldsymbol{x}_n$ 的密度依赖于 $\boldsymbol{x}_{n-1}$ 和 $\boldsymbol{v}_{n-1}$。如果 $\boldsymbol{v}_{n-1}$ 均值为零，那么 $\boldsymbol{x}_n$ 的平均值就是 $\boldsymbol{f}(\boldsymbol{x}_{n-1}, \boldsymbol{u}_n)$ 的平均值。

另外，由式(3－18)可知，z_n 可以被看作一个随机变量，它的当前值依赖于 $\boldsymbol{x}_n$。因此，式(3－1)定义了规范化的似然函数 $p(z_n \mid \boldsymbol{x}_n)/p(z_n)$。由式(3－1)可知，由于 $p(\boldsymbol{x})$ 和 $p(\boldsymbol{x} \mid z)$ 是概率密度，$p(z_n \mid \boldsymbol{x}_n)/p(z_n)$ 一定也是概率密度。因此，$p(z_n)$ 是 $p(z_n \mid \boldsymbol{x}_n)$ 的规范化函数。

图3－1给出了为得连续后验密度而进行的贝叶斯递推过程的一次迭代。递推过程从 $p(\boldsymbol{x}_0)$ 开始，与 $\boldsymbol{x}$ 相关的最先观测的概率密度函数。

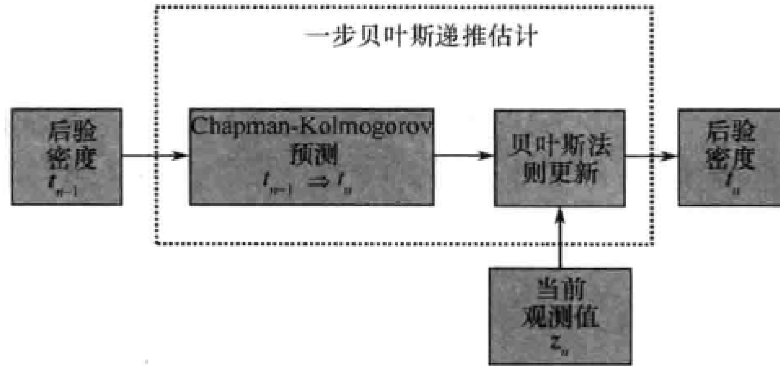


图3－1　一步递推贝叶斯后验密度估计过程

在理论上，式(3－17)中的马尔可夫过程可当作马尔可夫链 $\boldsymbol{x}_{1:n} = \{\boldsymbol{x}_1, \cdots, \boldsymbol{x}_n\}$ 的一部分，预测和更新方程，分别为式(3－24)和式(3－23)，写成马尔可夫链[2]形式。由于本书主要涉及跟踪应用，将马尔可夫链简化为一阶马尔可夫过程，推导过程中省略马尔可夫链符号，支持一步递推公式。对于动态模型高阶马尔可夫过程情况，如时间序列的数据分析，通常可以创建扩展状态矢量来得一阶马尔可夫过程。

3.4　基于状态均值和协方差的递推贝叶斯估计

这里引入一个符号 $\hat{\boldsymbol{x}}_{n|p}$，在本书的后续部分将使用这个符号。令 $\boldsymbol{x}_n$ 的估计以截至当前时刻 t_p 的所有观测为条件，则有

$$\hat{\boldsymbol{x}}_{n|p} = \mathcal{E}\{\boldsymbol{x}_n \mid z_{1:p}\} \triangleq \int_{\mathbf{R}^{n_x}} \boldsymbol{x}_n p(\boldsymbol{x}_n \mid z_{1:p}) \mathrm{d}\boldsymbol{x}_n \tag{3-25}$$

这样，$\boldsymbol{x}_n$ 的估计使用包括当前 t_n 时刻在内的所有观测值，那么

$$\hat{\boldsymbol{x}}_{n|n} = \mathcal{E}\{\boldsymbol{x}_n \mid z_{1:n}\} = \int_{\mathbf{R}^{n_x}} \boldsymbol{x}_n p(\boldsymbol{x}_n \mid z_{1:n}) \mathrm{d}\boldsymbol{x}_n \tag{3-26}$$

对 $\boldsymbol{x}_n$ 的预测是基于除当前观测之外的所有观测，即

$$\hat{\boldsymbol{x}}_{n|n-1} = \mathcal{E}\{\boldsymbol{x}_n \mid z_{1:n-1}\} = \int_{\mathbf{R}^{n_x}} \boldsymbol{x}_n p(\boldsymbol{x}_n \mid z_{1:n-1}) \mathrm{d}\boldsymbol{x}_n \tag{3-27}$$

对协方差矩阵做同样运算，得

$$\boldsymbol{P}_{n|n}^{xx} = \mathcal{E}\{(\boldsymbol{x}_n - \hat{\boldsymbol{x}}_{n|n})(\boldsymbol{x}_n - \hat{\boldsymbol{x}}_{n|n})^{\mathrm{T}} \mid z_{1:n}\}$$

$$= \int_{\mathbf{R}^{n_x}} (\boldsymbol{x}_n - \hat{\boldsymbol{x}}_{n|n})(\boldsymbol{x}_n - \hat{\boldsymbol{x}}_{n|n})^{\mathrm{T}} p(\boldsymbol{x}_n \mid \boldsymbol{z}_{1:n}) \mathrm{d}\boldsymbol{x}_n \quad (3-28)$$

协方差的预测形式如下

$$\boldsymbol{P}_{n|n}^{xx} = \mathcal{E}\{(\boldsymbol{x}_n - \hat{\boldsymbol{x}}_{n|n})(\boldsymbol{x}_n - \hat{\boldsymbol{x}}_{n|n})^{\mathrm{T}} \mid \boldsymbol{z}_{1:n}\} \quad (3-29)$$
$$= \int (\boldsymbol{x}_n - \boldsymbol{x}_{n|n})(\boldsymbol{x}_n - \hat{\boldsymbol{x}}_{n|n})^{\mathrm{T}} p(\boldsymbol{x}_n \mid \boldsymbol{z}_{1:n}) \mathrm{d}\boldsymbol{x}_n$$

3.4.1 状态矢量预测

将 Chapman - Kolmogorov 方程(3 - 24) 代入式(3 - 27) 和式(3 - 29)

得
$$\hat{\boldsymbol{x}}_{n|n-1} = \int_{\mathbf{R}^{n_x}} \int_{\mathbf{R}^{n_x}} \boldsymbol{x}_n p(\boldsymbol{x}_n \mid \boldsymbol{x}_{n-1}) p(\boldsymbol{x}_{n-1} \mid \boldsymbol{z}_{1:n-1}) \mathrm{d}\boldsymbol{x}_{n-1} \mathrm{d}\boldsymbol{x}_n \quad (3-30)$$

和
$$\boldsymbol{P}_{n|n-1}^{xx} = \int_{\mathbf{R}^{n_x}} \int_{\mathbf{R}^{n_x}} (\boldsymbol{x}_n - \hat{\boldsymbol{x}}_{n|n-1})(\boldsymbol{x}_n - \hat{\boldsymbol{x}}_{n|n-1})^{\mathrm{T}} p(\boldsymbol{x}_n \mid \boldsymbol{x}_{n-1}) \cdot$$
$$p(\boldsymbol{x}_{n-1} \mid \boldsymbol{z}_{1:n-1}) \mathrm{d}\boldsymbol{x}_{n-1} \mathrm{d}\boldsymbol{x}_n \quad (3-31)$$

根据系统动态方程(3 - 17),式(3 - 30) 可改写为

$$\hat{\boldsymbol{x}}_{n|n-1} = \int_{\mathbf{R}^{n_x}} \int_{\mathbf{R}^{n_x}} [\boldsymbol{f}_{n-1}(\boldsymbol{x}_{n-1}) + \boldsymbol{u}_n + \boldsymbol{v}_{n-1}] p(\boldsymbol{x}_n \mid \boldsymbol{x}_{n-1}) \cdot$$
$$p(\boldsymbol{x}_{n-1} \mid \boldsymbol{z}_{1:n-1}) \mathrm{d}\boldsymbol{x}_{n-1} \mathrm{d}\boldsymbol{x}_n$$
$$= \int_{\mathbf{R}^{n_x}} (\boldsymbol{f}_{n-1}(\boldsymbol{x}_{n-1}) + \boldsymbol{u}_n + \boldsymbol{v}_{n-1}] p(\boldsymbol{x}_{n-1} \mid \boldsymbol{z}_{1:n-1}) \mathrm{d}\boldsymbol{x}_{n-1} \cdot$$
$$\int_{\mathbf{R}^{n_x}} p(\boldsymbol{x}_n \mid \boldsymbol{x}_{n-1}) \mathrm{d}\boldsymbol{x}_n \quad (3-32)$$
$$= \int_{\mathbf{R}^{n_x}} \boldsymbol{f}_{n-1}(\boldsymbol{x}_{n-1}) p(\boldsymbol{x}_{n-1} \mid \boldsymbol{z}_{1:n-1}) \mathrm{d}\boldsymbol{x}_{n-1} + \boldsymbol{u}_n +$$
$$\int_{\mathbf{R}^{n_x}} \boldsymbol{v}_{n-1} p(\boldsymbol{x}_{n-1} \mid \boldsymbol{z}_{1:n-1}) \mathrm{d}\boldsymbol{x}_{n-1} \quad (3-33)$$

在式(3 - 32) 推导到式(3 - 33) 的过程中,密度函数的积分是涵盖整个支撑区域的密度,且其值总是为 1,因此积分值$\int p(\boldsymbol{x}_n \mid \boldsymbol{x}_{n-1}) \mathrm{d}\boldsymbol{x}_n = 1$。方程(3 - 33) 包括两个密度加权积分,除非在特殊情况下,否则一般无法用解析方法求解。如果噪声 $\boldsymbol{v}_{n-1}$ 均值为零,需要一个无偏估计器,第二个积分项简化为零。

类似地,式(3 - 31) 化简为

$$\boldsymbol{P}_{n|n-1}^{xx} = \int_{\mathbf{R}^{n_x}} [\boldsymbol{f}_{n-1}(\boldsymbol{x}_{n-1}) + \boldsymbol{u}_n - \hat{\boldsymbol{x}}_{n|n-1}][\boldsymbol{f}_{n-1}(\boldsymbol{x}_{n-1}) + \boldsymbol{u}_n - \hat{\boldsymbol{x}}_{n|n-1}]^{\mathrm{T}} \cdot$$
$$p(\boldsymbol{x}_{n-1} \mid \boldsymbol{z}_{1:n-1}) \mathrm{d}\boldsymbol{x}_{n-1} -$$
$$\int_{\mathbf{R}^{n_x}} [\boldsymbol{f}_{n-1}(\boldsymbol{x}_{n-1}) + \boldsymbol{u}_n - \hat{\boldsymbol{x}}_{n|n-1}] \boldsymbol{v}_{n-1}^{\mathrm{T}} p(\boldsymbol{x}_{n-1} \mid \boldsymbol{z}_{1:n-1}) \mathrm{d}\boldsymbol{x}_{n-1} -$$
$$(3-34)$$

$$\int_{\mathbf{R}^{n_x}} \boldsymbol{v}_{n-1}[\boldsymbol{f}_{n-1}(\boldsymbol{x}_{n-1}) + \boldsymbol{u}_n - \hat{\boldsymbol{x}}_{n|n-1}]^{\mathrm{T}} p(\boldsymbol{x}_{n-1} \mid z_{1:n-1})\mathrm{d}\boldsymbol{x}_{n-1} +$$

$$\int_{\mathbf{R}^{n_x}} \boldsymbol{v}_{n-1}\boldsymbol{v}_{n-1}^{\mathrm{T}} p(\boldsymbol{x}_{n-1} \mid z_{1:n-1})\mathrm{d}\boldsymbol{x}_{n-1}$$

定义噪声协方差为

$$\boldsymbol{Q} \triangleq \int_{\mathbf{R}^{n_x}} \boldsymbol{v}_{n-1}\boldsymbol{v}_{n-1}^{\mathrm{T}} p(\boldsymbol{x}_{n-1} \mid z_{1:n-1})\mathrm{d}\boldsymbol{x}_{n-1} \tag{3-35}$$

假定$\boldsymbol{f}_{n-1}(\boldsymbol{x}_{n-1},\boldsymbol{u}_n)$与$\boldsymbol{v}_{n-1}$不相关,式(3-34)中间两个积分化简为零,密度加权预测方差则变成

$$\boldsymbol{P}_{n|n-1}^{xx} = \int_{\mathbf{R}^{n_x}} [\boldsymbol{f}_{n-1}(\boldsymbol{x}_{n-1}) + \boldsymbol{u}_n - \hat{\boldsymbol{x}}_{n|n-1}][\boldsymbol{f}_{n-1}(\boldsymbol{x}_{n-1}) + \boldsymbol{u}_n - \boldsymbol{x}_{n|n-1}]^{\mathrm{T}} \cdot p(\boldsymbol{x}_{n-1} \mid z_{1:n-1})\mathrm{d}\boldsymbol{x}_{n-1} + \boldsymbol{Q} \tag{3-36}$$

因此,式(3-33)和式(3-36)状态矢量的一阶和二阶矩的预测估计是基于除最近一次观测外的所有观测为条件的。现在需要的是给定当前观测值z_n条件下的一种递归更新这些预测矩的方法。直接得到后验密度的前两阶矩,而无需评价式(3-26)和式(3-28)。

3.4.2 状态矢量更新

对于所有一般的密度,卡尔曼滤波器提出一种不采用贝叶斯法则[3]的滤波器。假定系统状态$\boldsymbol{x}_n$可以按次序得估计,仅更新一阶和二阶预测矩。具体而言,假定一阶矩的更新估计是线性依赖于当前观测值,即

$$\hat{\boldsymbol{x}}_{n|n} = \boldsymbol{A}z_n^{\mathrm{o}} + \boldsymbol{b} \tag{3-37}$$

式中:z_n^{o}为最新的含噪观测值;$\boldsymbol{A}$和$\boldsymbol{b}$为待确定的未知数。

根据当前观测值z_n^{o}(z_n^{o}为随机变量),随机变量$\boldsymbol{x}_n$最佳的线性估计(从最小均方误差意义上说)需满足如下条件:

(1)条件估计具有无偏性,即每个元素的估计误差式(3-9)均值为零,例如$\mathcal{E}\{\boldsymbol{\epsilon}_n^x \mid z_{1:n-1}\} = \boldsymbol{0} \in \mathbf{R}^{n_x}$;

(2)每个元素的估计误差与所有元素的观测值不相关,即$\mathcal{E}\{\boldsymbol{\epsilon}_n^x z_n^{\mathrm{oT}} \mid z_{1:n-1}\} = \boldsymbol{0} \in \mathbf{R}^{n_x \times n_x}$。

注意,上述两个条件都建立在估计误差之上,而不是在状态矢量或观测矢量之上。同时,后一个条件要求估计误差是与观测值正交。

根据式(3-9)中$\boldsymbol{\epsilon}_n^x$的定义和方程(3-37),运用条件(1),得

$$\begin{aligned}
\mathcal{E}\{\boldsymbol{\epsilon}_n^x \mid z_{1:n-1}\} &= \mathcal{E}\{\boldsymbol{x}_n - \hat{\boldsymbol{x}}_{n|n} \mid z_{1:n-1}\} \\
&= \mathcal{E}\{\boldsymbol{x}_n \mid z_{1:n-1}\} - (\boldsymbol{A}\,\mathcal{E}\{z_n^{\mathrm{o}} \mid z_{1:n-1}\} + \boldsymbol{b}) \\
&= \hat{\boldsymbol{x}}_{n|n-1} - \boldsymbol{A}\hat{z}_{n|n-1} - \boldsymbol{b} = \boldsymbol{0}
\end{aligned} \tag{3-38}$$

因此

$$b = -\hat{x}_{n|n-1} - A\hat{z}_{n|n-1} \tag{3-39}$$

由式(3 - 37)得

$$\hat{x}_{n|n} = \hat{x}_{n|n-1} + A(z_n^{\text{o}} - \hat{z}_{n|n-1}) \tag{3-40}$$

条件(2)是正交性条件,需要

$$\mathcal{E}\{\epsilon_n^x z_n^{\text{oT}} \mid z_{1:n-1}\} = \mathcal{E}\{[x_n - \hat{x}_{n|n-1} - A(z_n^{\text{o}} - \hat{z}_{n|n-1})]z_n^{\text{oT}}\} \tag{3-41}$$

注意,$\hat{x}_{n|n-1}$ 和 $\hat{z}_{n|n-1}$ 是以 $\hat{z}_{1:n-1}$ 为条件。因为$\mathcal{E}\{\epsilon_n^x\} = 0$, $\mathcal{E}\{\epsilon_n^x z_n^{\text{oT}}\} = \mathcal{E}\{\epsilon_n^x(z_n^{\text{o}} - \hat{z}_{n|n-1})^{\text{T}}\}$,因此式(3 - 41)可改写为

$$\begin{aligned}\mathcal{E}\{\epsilon_n^x z_n^{\text{oT}} \mid z_{1:n-1}\} &= \mathcal{E}\{[x_n - x_{n|n-1} - A(z_n^{\text{o}} - \hat{z}_{n|n-1})] \cdot \\ &\qquad (z_n^{\text{o}} - \hat{z}_{n|n-1})^{\text{T}} \mid z_{1:n-1}\} \\ &= P_{n|n-1}^{xz} - AP_{n|n-1}^{zz} = 0\end{aligned} \tag{3-42}$$

求解最后一个方程中的 A 值,得

$$A = K_n \triangleq P_{n|n-1}^{xz}[P_{n|n-1}^{zz}]^{-1} \tag{3-43}$$

用式(3 - 37)组合式(3 - 39)和式(3 - 43),得线性 MMSE 估计器,即

$$\hat{x}_{n|n} = \hat{x}_{n|n-1} + K_n(z_n^{\text{o}} - \hat{z}_{n|n-1}) \tag{3-44}$$

这里观测值的预测误差被称为新息(*innovations*),即

$$\mathcal{E}_n^z \triangleq \hat{z}_n^{\text{o}} - \hat{z}_{n|n-1} \tag{3-45}$$

更新后的协方差可以由下式得,即

$$\begin{aligned}P_{n|n}^{xx} &= \mathcal{E}\{\epsilon_n^x \epsilon_n^{x\text{T}} \mid z_{1:n-1}\} \\ &= \mathcal{E}\{(x_n - \hat{x}_{n|n})(x_n - \hat{x}_{n|n})^{\text{T}} \mid z_{1:n-1}\} \\ &= \mathcal{E}\{[x_n - \hat{x}_{n:n-1} - K_n(z_n^{\text{o}} - z_{n|n-1})] \cdot \\ &\qquad [x_n - \hat{x}_{n|n-1} - K_n(z_n^{\text{o}} - z_{n|n-1})]^{\text{T}} \mid z_{1:n-1}\}\end{aligned} \tag{3-46}$$

经过一些代数运算,式(3 - 46)可变成

$$P_{n|n}^{xx} = P_{n|n-1}^{xx} - K_n P_{n|n-1}^{zz} K_n^{\text{T}} \tag{3-47}$$

如上所述,式(3 - 44)和式(3 - 47)以及卡尔曼增益定义(3 - 43),都是使用最新的观测值来更新由式(3 - 33)和式(3 - 36)得的预测估计的卡尔曼滤波器更新方程。但是,上述方程需要额外的预测密度加权积分 $\hat{z}_{n|n-1}$、$P_{n|n-1}^{zz}$ 和 $P_{n|n-1}^{xz}$ 的计算。通常,这些预测值相关预测估计定义如下,即

$$\hat{z}_{n|n-1} \triangleq \varepsilon\{z_n \mid x_n, z_{1:n-1}\} \tag{3-48}$$

$$P_{n|n-1}^{zz} \triangleq \varepsilon\{(z_n - \hat{z}_{n|n-1})(z_n - \hat{z}_{n|n-1})^{\text{T}} \mid x_n, z_{1:n-1}\} \tag{3-49}$$

$$P_{n|n-1}^{xz} \triangleq \varepsilon\{(x_n - x_{n:n-1})(z_n - \hat{z}_{n|n-1})^{\text{T}} \mid z_{1:n-1}\} \tag{3-50}$$

用这些参数来推导式(3 - 33)和式(3 - 36),则要求有

$$\begin{aligned}\hat{z}_{n|n-1} &= \int_{\mathbf{R}^{n_x}} h_n(x_n)p(x_n \mid z_{1:n-1})\,\mathrm{d}x_n + \\ &\qquad \int_{\mathbf{R}^{n_x}} w_n p(x_n \mid z_{1:n-1})\,\mathrm{d}x_n\end{aligned} \tag{3-51}$$

$$P^{zz}_{n|n-1} = \int_{\mathbf{R}^{n_x}} [\boldsymbol{h}_n(\boldsymbol{x}_n) - \hat{z}_{n|n-1}][\boldsymbol{h}_n(\boldsymbol{x}_n) - \hat{z}_{n|n-1}]^{\mathrm{T}} \cdot p(\boldsymbol{x}_n \mid z_{1:n-1})\mathrm{d}\boldsymbol{x}_n + \boldsymbol{R} \quad (3-52)$$

$$P^{xz}_{n|n-1} = \int_{\mathbf{R}^{n_x}} [\boldsymbol{x}_n - \hat{\boldsymbol{x}}_{n|n-1}][\boldsymbol{h}_n(\boldsymbol{x}_n) - \hat{z}_{n|n-1}]^{\mathrm{T}} \cdot p(\boldsymbol{x}_n \mid z_{1:n-1})\mathrm{d}\boldsymbol{x}_n \quad (3-53)$$

$$\boldsymbol{R} \triangleq \int_{\mathbf{R}^{n_x}} \boldsymbol{w}_n \boldsymbol{w}_n^{\mathrm{T}} p(\boldsymbol{x}_n \mid z_{1:n-1})\mathrm{d}\boldsymbol{x}_n \quad (3-54)$$

式(3-52)中假定$\boldsymbol{h}_n(\boldsymbol{x}_n)$独立于$\mathbf{w}_n$,因此交叉项积分为零。通常假定$\mathbf{w}_n$均值为零,因此式(3-51)的第二个积分项积分值为零。

图3-2给出了任意分布下递推点估计过程的总体框图。

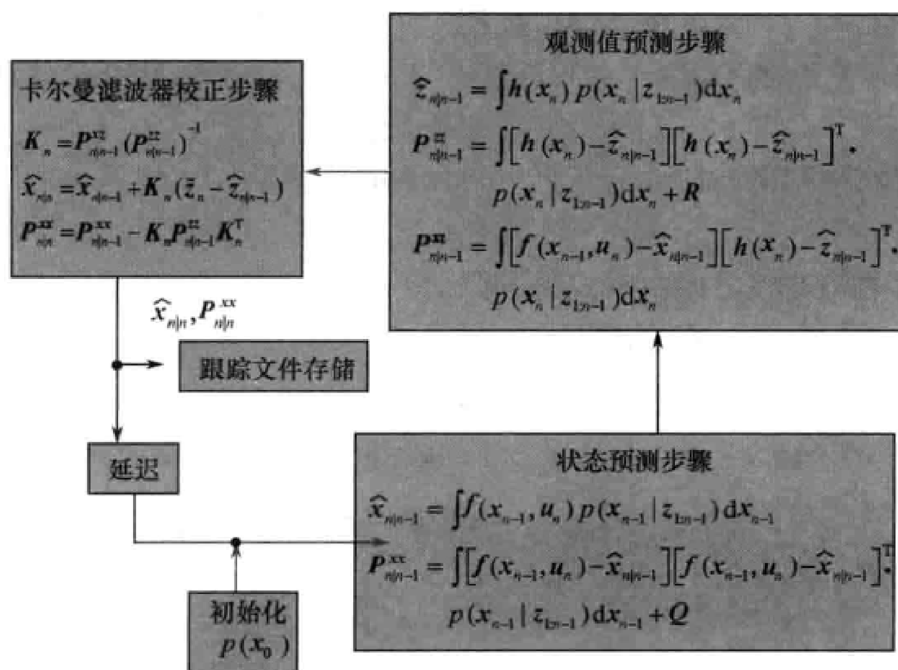


图3-2　递推点估计过程的总体框图

3.5　一般估计方法讨论

本章给出了贝叶斯估计的两个主要结果,在概率密度的形式和动态方程或观测方程的线性度上几乎没有任何限制。第一个结果(利用式(3-23)和式(3-24)得)建立前一个的后验概率$p(\boldsymbol{x}_{n-1} \mid z_{1:n-1})$和当前后验概率$p(\boldsymbol{x}_n \mid z_{1:n})$之间的递推关系,需要给出预测密度$p(\boldsymbol{x}_n \mid \boldsymbol{x}_{n-1})$和似然函数$p(z_n \mid \boldsymbol{x}_n)$。

在许多感兴趣的问题中,这些密度的递推蒙特卡洛评价提供了以所有观测值为条件,状态矢量可能的位置和传播的可视化洞察。本书第Ⅲ部分将扩展这些蒙特卡罗估计方法使其适用于非线性系统和非高斯密度情况(线性和高斯情

况作为一个子集被包含在内)。此外,提出的蒙特卡洛方法可以评价密度加权积分,其中 $\boldsymbol{x}_n$ 的蒙特卡洛采样可以用来生成密度函数,简化了采样点加权和的积分。后面这些方法被称为粒子滤波器。

本章提出的第二个估计过程是一个通用的方法,递推估计以所有观测值为条件的前两阶矩的状态矢量。这种方法使用密度加权积分预测估计 $\hat{\boldsymbol{x}}_{n|n-1}$、$\boldsymbol{P}_{n|n-1}^{xx}$、$\hat{\boldsymbol{z}}_{n|n-1}$、$\boldsymbol{P}_{n|n-1}^{zz}$ 和 $\boldsymbol{P}_{n|n-1}^{xz}$,分别由式(3 - 33)、式(3 - 36)、式(3 - 51)、式(3 - 52) 和式(3 - 53) 给出。然后是卡尔曼滤波更新方程式(3 - 43)、式(3 - 44) 和式(3 - 47)。

目前提出的方法都是通用的,对动态过程或观测过程的形式和概率密度的形式都没有限制。本书第 Ⅱ 部分致力于高斯密度下的点估计方法。在这个高斯假设基础上,导出本书中所有卡尔曼滤波器的变型。卡尔曼滤波器的变型主要分为三种类型,线性卡尔曼滤波器、扩展卡尔曼滤波器和 Sigma 点卡尔曼滤波器(数值计算高斯加权积分)。

参考文献

[1] Van Trees HL. *Detection, Estimation, and Modulation Theory, Part I: Detection, Estimation, and Linear Modulation Theory*. Wiley, 1968.

[2] Doucet A, Johansen AM. A Tutorial on Particle Filtering and Smoothing: Fifteen Years Later. Technical Report, Department of Statistics, University of British Columbia, 2008.

[3] Kalman RE. A new approach to linear filtering and prediction problems. *ASME J. Basic Eng*, 1960:35 -45.

第4章 实用案例：初步探讨

本书包括4个实用案例，用来举例说明第 Ⅱ 和第 Ⅲ 部分提出的跟踪滤波器的应用和性能。本章给出第一个案例，演示通过定向频率和测距（DIFAR）声纳浮标场进行船舶跟踪的问题。这是一个特别有趣的问题，当输入信号为高信噪比时，每个浮标的方位观测值的噪声是高斯分布的；而当输入信号为低信噪比时，噪声是非高斯分布的。这个案例可以检验低信噪比条件下跟踪滤波器的效用，此时的观测噪声是非高斯分布的。在第 Ⅱ 和第 Ⅲ 部分，利用这个案例说明多种跟踪滤波器都可以用在 DIFAR 问题上，并为每种估算方法提出实现细节，标绘出跟踪输出结果和真实航迹，最后评价所有跟踪算法的相关跟踪性能度量值。

本书第 Ⅳ 部分将讨论其它的案例分析，包括利用雷达观测数据跟踪三维空间中运动的物体；根据视频观测跟踪下降炸弹（摄影测量法）；以及在摄影测量法跟踪系统中使用多种跟踪滤波器数据融合。这些案例分析的范围从相对简单的 DIFAR 纯方位跟踪问题，到非常复杂的需要高维状态矢量、观测矢量和高数据速率的问题。对于某些问题，所有的跟踪滤波器都能得基本相同的结果；而对其它一些问题，有些跟踪滤波器则无法使用，因为这些跟踪滤波器无法满足问题的内在要求。

4.1 仿真、估计与评价过程

为特定的跟踪问题场景编写软件代码，是一项很艰巨的任务。在图4－1中，给出了必须解决的问题过程框图。首先需要建立一个场景模拟器，生成待估计状态的真实数据；然后将这个真实数据集转换为一组真实观测值；最后添加噪声。

例如，在三维笛卡儿坐标中跟踪一辆车。首先要生成车辆在笛卡儿坐标中的真实位置和速度分量（也称作真实轨迹）。如果观测值来自于雷达，笛卡儿坐标真实轨迹必须从笛卡儿坐标变换到球坐标，生成不含噪声的球坐标真实轨迹，包

括测距、方位、高度和分别的速度。然后，把具体的雷达观测值（如测距、方位、高度、和／或多普勒距离变化率）从球坐标真实值中分离出来。在使用跟踪滤波器处理真实雷达获得的数据时，最后的仿真步骤可以省略。仿真步骤主要是用来测试和调试估计算法的执行情况，比较评估跟踪算法的性能。

在第二个方框中，通常是测试蒙特卡洛运动跟踪算法，并在观测值中加入噪声，将含噪观测值作为跟踪滤波器的输入。所研究的跟踪滤波器生成以时间为变量的跟踪状态矢量，以及跟踪状态误差协方差估计函数。



图 4－1　跟踪算法的研究与评估

最后一个方框中对跟踪算法的性能进行了评估。对于仿真观测值，可以将真实跟踪数据和跟踪估计用在一系列蒙特卡洛运行中，生成均方根（RMS）误差并将均方根误差与 Cramer－Rao 下界（CRLB）进行比较。当使用实际系统数据时，通常无法得到真实航迹，可以使用估计的跟踪误差协方差来生成跟踪估计的误差椭圆。

整个仿真中的每个流程框可以独立开发，因为在多数情况下，这些流程相对具体跟踪算法具有不变性。

下节将讨论 DIFAR 跟踪场景仿真器的算法基础，并根据具体问题分析算法的细节。对于第 Ⅱ 和第 Ⅲ 部分提出的每种航迹估计方法，作为每种算法的具体应用实例，将给出 DIFAR 问题的航迹估计结果。在 14.5 节对各种高斯估计方法处理 DIFAR 问题的性能进行比较，非高斯噪声情况下粒子滤波器估计方法的性能比较将在本书第 Ⅲ 部分末尾加以阐述。

4.2　利用 DIFAR 浮标场进行匀速运动目标跟踪的场景仿真

这里的 DIFAR 场景包括匀速通过 DIFAR 浮标场的船舶，使用的坐标系是二维笛卡儿坐标系，以浮标场的中心为坐标原点，x 轴指向东，y 轴指向北。

4.2.1　船舶动态模型

用状态矢量 $\boldsymbol{x}(t)$ 来描述 t 时刻船舶的运动，矢量由二维坐标下的位置和速度分量构成

$$\boldsymbol{x}(t) = [r^x(t) \quad v^x \quad r^y(t) \quad v^y]^{\mathrm{T}} \tag{4-1}$$

其中，$\{r^x(t), r^y(t)\}$ 和 $\{v^x, v^y\}$ 分别为笛卡儿坐标下的位置和速度分量。这里采

用了匀速运动模型，速度分量独立于时间。假定场景开始时船舶的位置在 $\{r^x(t_0), r^y(t_0)\}$，那么在时间 t 时，场景中船舶的位置分量可写成

$$r^x(t) = r^x(t_0) + v^x t \tag{4-2}$$

$$r^y(t) = r^y(t_0) + v^y t \tag{4-3}$$

用矢量矩阵表示，可改写为

$$\boldsymbol{x}(t) = \boldsymbol{F}_t \boldsymbol{x}(t_0) \tag{4-4}$$

$$\boldsymbol{F}_t = \begin{bmatrix} \boldsymbol{F}_{t,1} & \boldsymbol{0}_2 \\ \boldsymbol{0}_2 & \boldsymbol{F}_{t,1} \end{bmatrix} \tag{4-5}$$

$$\boldsymbol{F}_{t,1} = \begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix} \tag{4-6}$$

4.2.2 多浮标观测模型

假设有 M 个浮标随机分布在以原点为圆心、半径为 $R_{\max}$ 的区域内，第 m 个浮标位于固定位置 $\{x_m, y_m\}$。当然，在实际浮标场跟踪运动中，由于水流漂移和波浪作用，这种在跟踪估计时间帧中浮标位置固定的假设条件不可能实现。为便于算法证明，这里假定浮标位置固定。

将方位定义为相对正北（y 轴方向）的一个顺时针角度，可以将 t 时刻第 m 个浮标与目标之间的方位角写作

$$\theta_m(t) = \tan^{-1}\left(\frac{r^x(t) - x_m}{r^y(t) - y_m}\right) \tag{4-7}$$

式中：$-\pi \leqslant \theta_m(t) \leqslant \pi$。

4.2.3 场景属性

场景生成的开始两步为：

（1）利用式（4－2）和式（4－3）生成均匀间隔时间集合 $\{t_n; n = 0, 1, \cdots, N\}$ 上的真实航迹；

（2）进行笛卡儿坐标至极坐标的变换，用式（4－7）对场景中的时间帧生成没有噪声的观测值。

图 4－2 给出在以（0,0）为中心的浮标场中，目标船舶跟踪的几何原理。

图 4－3 给出船舶的真实航迹、真实方位与两个 DIFAR 传感器之间的关系。图 4－3（a）显示了一小部分船舶从左下角向右上角运动的真实航迹，还给出了原点周围随机分布的 15 个浮标的浮标场的示例。图 4－3（b）给出了船舶与左图中分别标记为 * 和 x 的两个浮标之间的方位角。图中浮标相对于坐标原点的距离和方位角的分布服从下式

$$\rho_m = R_{\max} * u_m, m = 1, \cdots, M$$

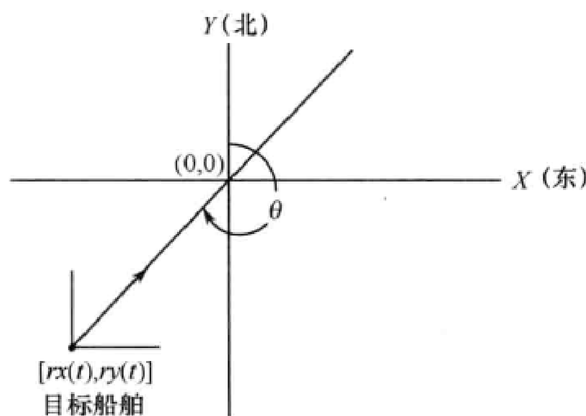


图 4 - 2　DIFAR 浮标几何原理

$$\vartheta_m \sim U(-\pi,\pi), m = 1,\cdots,M$$

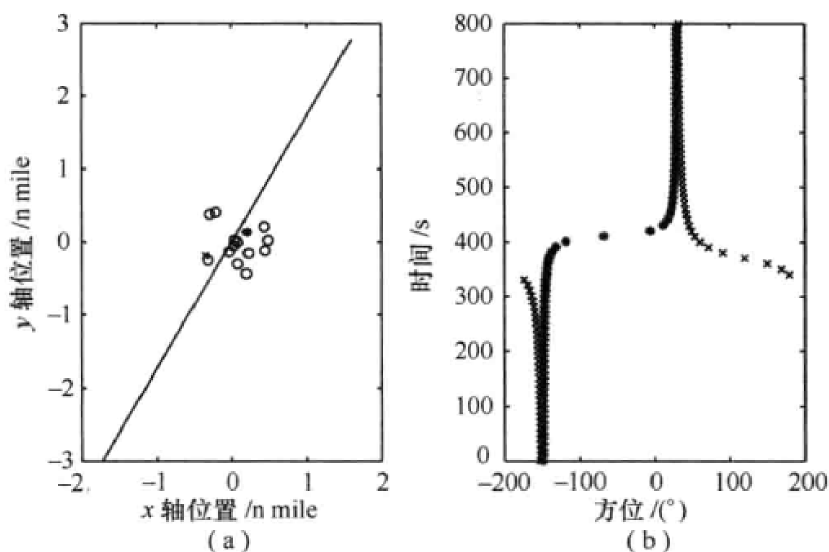


图 4 - 3　利用 DIFAR 浮标场和两个浮标的联合方位角进行船舶跟踪

其中 R_{max} 设为152.4m(500ft)[$u_m \sim U(a,b)$ 表示 u_m 是从区间 (a,b) 中均匀分布得]。将极坐标表示的浮标位置集合 $\{\rho_m,\vartheta_m; m = 1,\cdots,M\}$ 用下式变换为笛卡儿坐标集合 $\{x_m, y_m; m = 1,\cdots,M\}$，即

$$x_m = \rho_m \sin \vartheta_m \tag{4-8}$$

$$y_m = \rho_m \cos \vartheta_m \tag{4-9}$$

图4 - 3(b) 显示了方位角相关的两个标记浮标，一个浮标在船舶航迹的一侧。由图所示，浮标方位为东(右侧) 时，船舶航迹为 0°；当浮标方位为西时，船舶航迹为 - 180°，然后跳转到 180°。

方位为西的浮标经过了所谓的相位卷褶，相位角跳转 360°。第 6 章将给出在所有跟踪算法中相位卷褶所带来的问题，为使跟踪器能正常工作，必须对其加以修正。类似的相位卷褶问题也发生在许多不同领域中的估计方法的应用上，为得有效的结果，必须加以解决。列表4 - 1 中给出了一个 MATLAB 子程序，可以用来生成船舶在笛卡儿坐标和球坐标下的真实航迹，得一组随机的浮标位置。

列表 4 - 1 生成真实航迹及浮标位置

```
function [x_true,r_true,f,s,SNR,BT,SensorLocations,
R,theta_sigma,t] = DifarScenarioGenerator(Stream,...
N,M,Rmax,del_t,speed,heading,y_init)

global deg2rad hours2seconds
measurement_sigma = 3 * deg2rad;
theta_sigma = 1;

t = del_t * (1:N);
speed = speed/hours2seconds;
heading = heading * deg2rad;
vx_init = speed * sin(heading);
vy_init = speed * cos(heading);

x_init = y_init * (vx_init/vy_init);
x_true_init = [x_init;vx_init;y_init;vy_init];

r_m = Rmax * rand(Stream,M,1);
theta_m = 2 * pi * (rand(Stream,M,1) - 0.5);
xx = r_m. * sin(theta_m) + Sensor_Offset(1);
yy = r_m. * cos(theta_m) + Sensor_Offset(2);
Sensor_Locations = [xx yy];

R = diag(measurement_sigma - 2 * ones(M,1),0);

[x_true,f,s,SNR,BT] = ...
Constant_Velocity_Kinematics_Difar(Stream,...
N,t,x_true_init);

[r_true] = Cart2Polar(x_true,Sensor_Locations,N,M);
```

本书第 Ⅱ 和第 Ⅲ 部分将提出一些基于一组含噪观测值进行船舶航迹顺序(时间上)估计的算法。对于每种估计方法,利用 DIFAR 观测值作为滤波跟踪估计的输出结果。

4.3 DIFAR 浮标信号处理

DIFAR 传感器(一种矢量传感器)是一个被动接收器,如图 4 - 4 所示,对远处船舶辐射的宽带信号进行运算,将其变换为含噪船舶方位信号。DIFAR 浮标

由 3 个独立的传感器组成：全向水听器、东西偶极子水听器和南北偶极子水听器。两个偶极子水听器可能不按照规定的方向排列，但将其输出进行电子混合，可以得两个偶极子在正确方向的输出。后面不会进一步讨论这个重定位过程，因为它与跟踪问题无关，通常假定每个通道的输出都是沿笛卡儿坐标轴方向的。

在每个模拟 DIFAR 传感器输出中，需要生成含噪的方位观测值以实现仿真。如图 4－4 所示，对于第 m 个传感器，每个通道输出的真实信号如下式所示，即

$$z_{NS}(t) = s(t)\cos\theta + n_{NS}(t) \tag{4-10}$$

$$z_{EW}(t) = s(t)\sin\theta + n_{EW}(t) \tag{4-11}$$

$$z_{0}(t) = s(t) + n_{0}(t) \tag{4-12}$$

式中：$s(t)$ 为 DIFAR 浮标的输入信号。

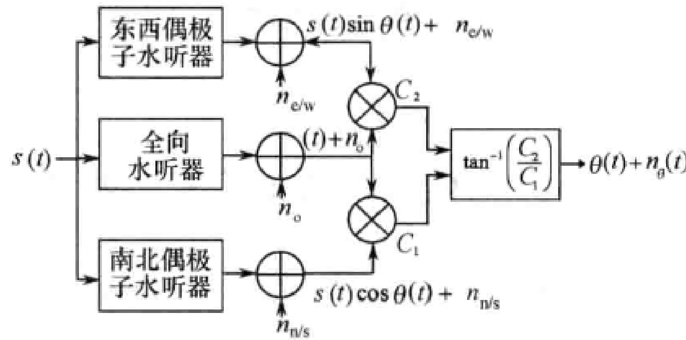


图 4－4　DIFAR 传感器信号处理

互相关系数 $C_1(t)$ 和 $C_1(t)$ 定义为

$$\begin{aligned} C_1(t) &= z_{NS}(t) \times z_0(t) \\ &= s^2(t)\cos\theta + s(t)n_0(t)\cos\theta + \\ &\quad s(t)n_{NS}(t) + n_{NS}(t)n_0(t) \end{aligned} \tag{4-13}$$

$$\begin{aligned} C_2(t) &= z_{EW}(t) \times z_0(t) \\ &= s^2(t)\sin\theta + s(t)n_0(t)\sin\theta + \\ &\quad s(t)n_{EW}(t) + n_{EW}(t)n_0(t) \end{aligned} \tag{4-14}$$

假定 θ 随时间变化缓慢，$C_1(t)$ 和 $C_2(t)$ 的期望值可写为

$$\mathcal{E}\{C_1(t)\} = \mathcal{E}\{s^2(t)\}\cos\theta \tag{4-15}$$

$$\mathcal{E}\{C_2(t)\} = \mathcal{E}\{s^2(t)\}\sin\theta \tag{4-16}$$

假定信号和噪声分量是零均值高斯过程，与传感器和时间无关。这样，由式(4－17)得方位角为

$$\theta(t) = \tan^{-1}\frac{\mathcal{E}\{C_2(t)\}}{\mathcal{E}\{C_1(t)\}} \tag{4-17}$$

为得$\mathcal{E}\{C_1(t)\}$ 和$\mathcal{E}\{C_2(t)\}$ 的估计值，对 $z_{NS}(t)$、$z_{EW}(t)$ 和 $z_0(t)$ 进行采样，

把时间划分成长度为 T 的段,令

$$\tau_q = t_n + (i-1)\delta_t - \frac{N}{2}\delta_t, i = 1,\cdots,N; n = 1,2,\cdots,\infty; q = 1,2,\cdots,\infty$$

$$\delta = \mathrm{T/N} \tag{4-18}$$

例如,可写成

$$\tau_1 = t_1 - \frac{N}{2}\delta_t$$

$$\tau_2 = t_1 + \delta_t - \frac{N}{2}\delta_t$$

$$\vdots$$

$$\tau_{\mathrm{N}} = t_1 + (\mathrm{N}-1)\delta_t - \frac{N}{2}\delta_t = t_1 + \left(\frac{N}{2}-1\right)\delta_t$$

$$\tau_{\mathrm{N+1}} = t_2 - \frac{N}{2}\delta_t \tag{4-19}$$

在 t_n 时刻,$\varepsilon\{C_1(t_n)\}$ 和 $\varepsilon\{C_2(t_n)\}$ 可由式(4-20)和式(4-21)估计得。

$$\varepsilon\{C_1(t_n)\} \simeq \int_{t_n-\frac{N}{2}\delta_t}^{t_n+\frac{N}{2}\delta_t} z_{\mathrm{NS}}(t)z_{\mathrm{O}}(t)\,\mathrm{d}t \tag{4-20}$$

$$\varepsilon\{C_2(t_n)\} \simeq \int_{t_n-\frac{N}{2}\delta_t}^{t_n+\frac{N}{2}\delta_t} z_{\mathrm{EW}}(t)z_{\mathrm{O}}(t)\,\mathrm{d}t \tag{4-21}$$

由 Weiner - Khinchine 定理[1],式(4-20)和式(4-21)可写为

$$\varepsilon\{C_1(t_n)\} = \int_{-\infty}^{\infty} Y_{\mathrm{NS}}(t_n,f_k)Y_{\mathrm{O}}^*(t_n,f_k)\,\mathrm{d}f \tag{4-22}$$

$$\varepsilon\{C_2(t_n)\} = \int_{-\infty}^{\infty} Y_{\mathrm{EW}}(t_n,f_k)Y_{\mathrm{O}}^*(t_n,f_k)\,\mathrm{d}f \tag{4-23}$$

这里 * 表示复共轭运算,且有

$$Y_{\mathrm{NS}}(t_n,f_k) = S(t_n,f_k)\cos\theta + N_{\mathrm{NS}}(t_n,f_k) \tag{4-24}$$

$$Y_{\mathrm{EW}}(t_n,f_k) = S(t_n,f_k)\sin\theta + N_{\mathrm{EW}}(t_n,f_k) \tag{4-25}$$

$$Y_{\mathrm{O}}(t_n,f_k) = S(t_n,f_k) + N_{\mathrm{O}}(t_n,f_k) \tag{4-26}$$

对于采样数据,时域方程(4-20)和(4-21)变成

$$\varepsilon\{C_1(t_n)\} \simeq \frac{1}{T}\sum_{i=1}^{\mathrm{N}} z_{\mathrm{NS}}(a_i\delta_t)z_{\mathrm{O}}(a_i\delta_t) \tag{4-27}$$

$$\varepsilon\{C_2(t_n)\} \simeq \frac{1}{T}\sum_{i=1}^{N} z_{\mathrm{EW}}(a_i\delta_t)z_{\mathrm{O}}(a_i\delta_t) \tag{4-28}$$

其中

$$a_i \triangleq i - \frac{N+2}{2} \tag{4-29}$$

在频域上，对集合 $\{z_{NS}(t), z_{EW}(t), z_O(t)\}$ 中的 N 个采样点进行快速傅里叶变换（FFT），得频域采样集合 $\{Y_{NS}(t_n, f_k), Y_{EW}(t_n, f_k), Y_O(t_n, f_k)\}$，进而导出

$$\varepsilon\{C_1(t_n)\} \simeq \frac{\delta_f}{N/2}\sum_{k=1}^{N/2} Y_{NS}(t_n, f_k) Y_O^*(t_n, f_k) \tag{4-30}$$

$$\varepsilon\{C_2(t_n)\} \simeq \frac{\delta_f}{N/2}\sum_{k=1}^{N/2} Y_{EW}(t_n, f_k) Y_O^*(t_n, f_k) \tag{4-31}$$

表 4-1 给出了根据船舶真实航迹生成每个 DIFAR 传感器含噪方位观测值的流程。列表 4-2 中给出生成含噪宽带信号的 MATLAB 程序片段。列表 4-3 为实现表 4-1 所列流程的 MATLAB 子程序。

表 4-1　生成含噪 DIFAR 观测矢量的流程

第 1 步	生成理想方位	$\theta_m(t) = \tan^{-1}\left(\frac{r^x(t) - x_m}{r^y(t) - y_m}\right)$
第 2 步	生成每个通道的复合随机噪声	$N_{NS}(t_n, f_k), N_{EW}(t_n, f_k), N_O(t_n, f_k)$
第 3 步	生成复合随机信号	$S(t_n, f_k)$
第 4 步	计算 DIFAR 通道输出	$Y_{NS,m}(t_n, f_k) = S(t_n, f_k)\cos\theta_m + N_{NS}(t_n, f_k)$ $Y_{EW,m}(t_n, f_k) = S(t_n, f_k)\sin\theta_m + N_{EW}(t_n, f_k)$ $Y_O(t_n, f_k) = S(t_n, f_k) + N_O(t_n, f_k)$
第 5 步	计算卷积	$C_{1,m}(t_n) = Y_{NS}(t_n, f_k) Y_O^*(t_n, f_k)$ $C_{2,m}(t_n) = Y_{EW}(t_n, f_k) Y_O^*(t_n, f_k)$
第 6 步	频带累加	$\varepsilon\{C_1, m(t_n)\} = \frac{\delta_f}{N/2}\sum_{k=1}^{N/2} C_{1,m}(t_n)$ $\varepsilon\{C_{2,m}(t_n)\} = \frac{\delta_f}{N/2}\sum_{k=1}^{N/2} C_{2,m}(t_n)$
第 7 步	计算含噪方位	$\theta_{n,m} = \tan^{-1}\left[\frac{\varepsilon\{C_{2,m}(t_n)\}}{\varepsilon\{C_1, m(t_n)\}}\right]$

列表 4-2 在 K 个频率段中生成 N 个含噪信号采样的 MATLAB 代码片段

```
f = 300:15:400;
f_target = [10 20 50 100 200 400 1000 4000]';

SPL_target = [20;20;20;20;20;20;20;20];

SPL = interp1q(log10(f_target),SPL_target,log10(f'));
```

```
spl = repmat(10.^(SPL/10),1,N);
SNR = spl(1);
K = length(f);
BT = (f(K) - f(1))*(t(2) - t(1));

SPL = interp1q(log10(f_target),SPL_target,log10(f'));
spl = repmat(10.^(SPL/10),1,N);
SNR = spl(1);
K = length(f);

BT = (f(K) - f(1))*(t(2) - t(1));

s = sqrt(spl/2).*(randn(Stream,K,N) + ...
1i*randn(Stream,K,N));

K = length(f);
BT = (f(K) - f(1))*(t(2) - t(1));
```

列表4－3 按照图4－1中步骤生成含噪方位测量值的MATLAB子程序

```
function [bearing_measurement] = ...
First_Sensor_Measurement_Generator_Difar(f,s,...
theta_sigma,theta_true,N,M,Stream)

K = length(f);
del_f = f(2) - f(1);

NS_ms = sqrt(theta_sigma)*(randn(Stream,K,N,M) + ...
1i*randn(Stream,K,N,M));
EW_ms = sqrt(theta_sigma)*(randn(Stream,K,N,M) + ...
1i*randn(Stream,K,N,M));
O_ms = sqrt(theta_sigma)*(randn(Stream,K,N,M) + ...
1i*randn(Stream,K,N,M));

Y_NS = repmat(s,[1,1,M]).*...
repmat(cos(theta_true),[K,1,1]) + NS_ms;
Y_EW = repmat(s,[1,1,M]).*...
repmat(sin(theta_true),[K,1,1]) + EW_ms;
Y_O = repmat(s,[1,1,M]) + O_ms;

C1 = real(Y_NS.*conj(Y_O));
```

```
C2 = real(Y_EW.*conj(Y_O));

E_C1 = (2*del_f/N)*permute(sum(C1),[2 3 1]);
E_C2 = (2*del_f/N)*permute(sum(C2),[2 3 1]);

bearing_measurement = permute(atan2(E_C2,E_C1),[2 1]);
```

必须注意,表4-1中第7步包含的反正切运算里的比值是随机数的比值。众所周知,两个独立零均值变量比值的分布为柯西分布。然而,如果两个变量均值不为零,其比值分布就要复杂得多,这在下节中将会看到。此外,对于 DIFAR 方位,反正切运算的分布更为复杂,所以不能期望 DIFAR 方位观测值符合高斯分布。但是,可以期望分布是 DIFAR 浮标输入信号信噪比的函数。为说明这一点,图4-5给出了4个不同的信号信噪比(DIFAR 传感器的输入)条件下,浮标在100次蒙特卡洛运行中特定时刻的方位观测值。很明显,随着信噪比的增加,分布曲线变得更加突兀,近似于高斯分布;而随着信噪比的降低,方位分布近似于[-180°,180°]区间的均匀分布。

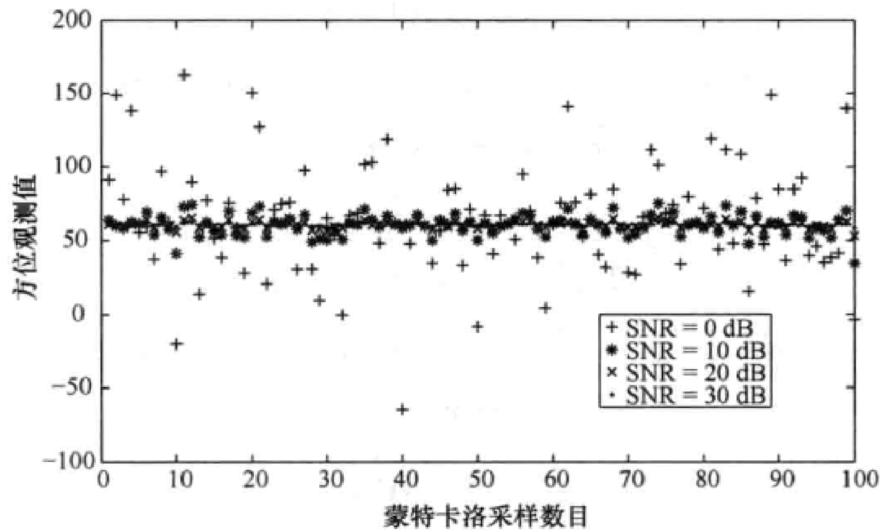


图4-5 4种信噪比下、100次蒙特卡洛运行得的 DIFAR 浮标的方位观测值的分布

4.4 DIFAR 似然函数

如果假设 DIFAR 传感器似然函数独立于传感器,以目标船舶的位置为条件的观测矢量似然函数可由联合密度给出,即

$$p(\boldsymbol{z}_n \mid \boldsymbol{x}_n) = p(\boldsymbol{\theta}_n \mid \boldsymbol{x}_n) = \prod_{m=1}^{M} p(\boldsymbol{\theta}_{n,m} \mid \boldsymbol{x}_n) \tag{4-32}$$

其中 $\boldsymbol{\theta}_n = [\theta_{n,1},\cdots,\theta_{n,M}]^{\mathrm{T}}$ 且

$$\theta_{n,m} = \tan^{-1}\left(\frac{r_n^x - x_m}{r_n^y - y_m}\right) \tag{4-33}$$

虽然它本身是一个有趣的问题,但 DIFAR 浮标似然函数的完整推导并不是本书的重点,感兴趣的读者可以在附录的在线参考文献[2] 中找到推导过程。因为在本书第 Ⅲ 部分需要用到,这里给出似然密度结果及其性质。

为描述清晰,省略时间索引,令 $z_m \triangleq \tan\theta_m$,根据参考文献[2] 的附录 A,可以有

$$p(\theta_m \mid \boldsymbol{x}_n) = [1 + (z_m)^2]p(z_m \mid \theta_{n,m}) \tag{4-34}$$

$$p(z_m \mid \theta_m) = A\left\{\frac{\sqrt{(1-r^2)}}{\pi}\mathrm{e}^{-\alpha} + \frac{\delta\beta^{1/2}}{\sqrt{2\pi}}\cdot\left[\Phi\left(\frac{-\delta\beta^{1/2}}{\sqrt{(1-r^2)}}\right) - \Phi\left(\frac{\delta\beta^{1/2}}{\sqrt{(1-r^2)}}\right)\right]\mathrm{e}^{-\gamma}\right\} \tag{4-35}$$

$$A \triangleq \frac{\tilde{\sigma}_1\tilde{\sigma}_2}{\tilde{\sigma}_2^2 - 2r\tilde{\sigma}_1\tilde{\sigma}_2 z_m + \tilde{\sigma}_1^2 z_m^2} \tag{4-36}$$

$$\alpha \triangleq \frac{\tilde{\sigma}_2^2\eta_1^2 - 2r\tilde{\sigma}_1\tilde{\sigma}_2\tilde{\eta}_1\tilde{\eta}_2 + \tilde{\sigma}_1^2\tilde{\eta}_2^2}{2\tilde{\sigma}_1^2\tilde{\sigma}_2^2(1-r^2)} \tag{4-37}$$

$$\beta \triangleq \frac{[(\tilde{\sigma}_2^2\tilde{\eta}_1 - r\tilde{\sigma}_1\tilde{\sigma}_2\tilde{\eta}_2) + (\tilde{\sigma}_1^2\tilde{\eta}_2 - r\tilde{\sigma}_1\tilde{\sigma}_2\tilde{\eta}_1)z_m]^2}{\tilde{\sigma}_1^2\tilde{\sigma}_2^2(\tilde{\sigma}_2^2 - 2r\tilde{\sigma}_1\tilde{\sigma}_2 z_m + \tilde{\sigma}_1^2 z_m^2)} \tag{4-38}$$

$$\gamma \triangleq \frac{(\tilde{\eta}_2 - \tilde{\eta}_1 z_m)^2}{2(\tilde{\sigma}_2^2 - 2r\tilde{\sigma}_1\tilde{\sigma}_2 z_m + \tilde{\sigma}_1^2 z_m^2)} \tag{4-39}$$

$$\delta \triangleq \operatorname{sign}[(\tilde{\sigma}_2^2\tilde{\eta}_1 - r\tilde{\sigma}_1\tilde{\sigma}_2\tilde{\eta}_2) + (\tilde{\sigma}_1^2\,\tilde{\eta}_2 - r\tilde{\sigma}_1\tilde{\sigma}_2\tilde{\eta}_1)z_m] \tag{4-40}$$

$$\tilde{\eta}_1 \triangleq \sqrt{2\mathrm{BTSNR}}\cos\theta_m^{(i)} \tag{4-41}$$

$$\tilde{\eta}_2 \triangleq \sqrt{2\mathrm{BTSNR}}\sin\theta_m^{(i)} \tag{4-42}$$

$$\tilde{\sigma}_1^2 \triangleq (2\mathrm{SNR}+1)\mathrm{SNR}\cos^2\theta_m^{(i)} + \rho(\mathrm{SNR}+1) \tag{4-43}$$

$$\tilde{\sigma}_2^2 \triangleq (2\mathrm{SNR}+1)\mathrm{SNR}\sin^2\theta_m^{(i)} + \rho(\mathrm{SNR}+1) \tag{4-44}$$

$$r \triangleq \frac{(2\mathrm{SNR}+1)\mathrm{SNR}\cos\theta_m^{(i)}\sin\theta_m^{(i)}}{\tilde{\sigma}_1\tilde{\sigma}_2} \tag{4-45}$$

$$\Phi(x) \triangleq \frac{1}{\sqrt{2\pi}}\int_x^{\infty}\mathrm{e}^{-u^2/2}\mathrm{d}u \tag{4-46}$$

信噪比是单极通道下信号与噪声的比值,BT 为车辆频率特征的时宽带宽积,ρ 为适用于偶极通道的噪声增益(减损) 因子 $\Phi(x)$ 与误差函数相关。对于二维各项同性和三维各项同性噪声,ρ 的取值分别为 1/2 或 1/3[3]。

注意,式(4 - 34) 给出的似然函数实际是条件似然概率 $p(\theta \mid \theta_0)$,其中 θ_0 为由式(4 - 7) 得的相对目标船舶的真实方位,θ 代表计算似然概率用到的方位值

集合。图 4 - 6 给出带宽时宽积(BT) 为 10 且信噪比取值范围为(- 10 ≤ SNR ≤ 10) 时标绘出的($p(\theta \mid \theta_0)$) 值。很容易看出,对于低信噪比情况,似然密度远离高斯分布,这是由于它过高的拖尾,还有它仅在 $-\pi \leqslant \theta \leqslant \pi$ 区间范围有定义。

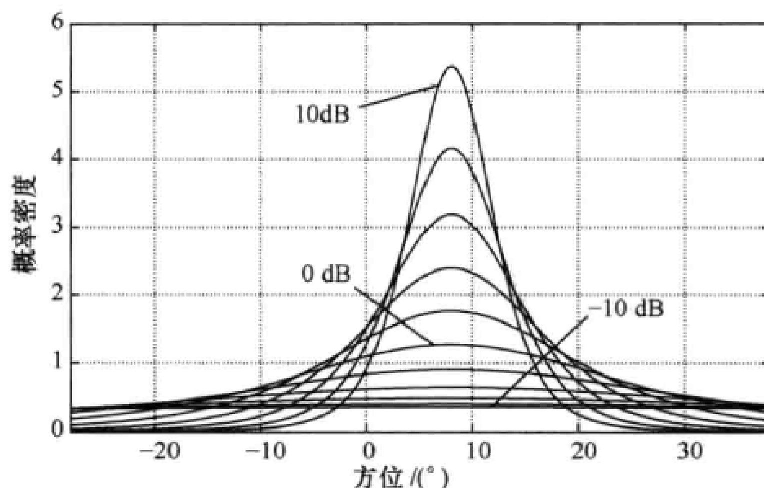


图 4 - 6 DIFAR 似然密度(BT 取值为 10,信噪比从 - 10 到 + 10dB)

参考文献

[1] Marple SL. *Digital Spectral Analysis with Applications.* Prentice - Hall: Englewood Cliffs, NJ, 1987.

[2] Haug AJ. A Tutorial on Bayesian Estimation and Tracking Techniques Applicable to Nonlinear and Non - Gaussian Processes. MITRE Technical Report MTR 05W0000004. 2005.

[3] Maranda BH. The statistical accuracy of an arctangent bearing estimator. *OCEANS* 2003, 4: 2127 - 2132.

BAYESIAN

第Ⅱ部分

高斯假设:卡尔曼滤波估计器

第5章
高斯噪声:高斯加权分布的多维积分

第3章结尾部分采用一般概率密度函数导出了贝叶斯点估计方程。本章假定高斯过程由动态方程和观测方程构成。在这种假设条件下,点估计方程中包含的所有的分布函数都是高斯分布。众所周知,高斯密度用其前两阶矩就可以完整表示。因此,对于高斯过程来说,前两阶矩的递归传播估计是最优的估计方法。这一部分的主题包括多种算法的推导,用来求解第3章提出的在高斯密度情况下的加权密度预测点估计。

在第3.4节,提出了一组卡尔曼滤波更新方程,对过程是否为线性以及过程的密度没有特殊要求。假设 $\boldsymbol{x}_n$(其后验点估计写成 $\hat{\boldsymbol{x}}_{n|n}$)一定为最新观测 $\boldsymbol{z}_n^o$ 的线性函数。点估计器的条件是:估计误差均值为零且与观测矢量互不相关(译者注:此处原文有误)。在后面章节中将介绍如何由高斯分布得卡尔曼滤波更新方程,高斯分布能满足所有基本假设和要求。

以第3章提出的一般形式的贝叶斯点估计方程为基础,第5.3节主要说明将预测估计方程简化为高斯加权积分。在后面章节,针对各种假设以及不同的动态函数 $\boldsymbol{f}_n(\boldsymbol{x}_n)$ 和观测函数 $\boldsymbol{h}_n(\boldsymbol{x}_n)$ 的逼近,用数值方法求解卡尔曼滤波器预测高斯加权积分。根据求解方法的不同,将卡尔曼滤波器分为以下4类:

(1) 线性卡尔曼滤波器。当动态方程和观测方程均为线性,且所有密度均是高斯分布,积分可以直接求解,这就是线性卡尔曼滤波器(the Linear Kalman Filter,LKF)。

(2) 解析线性化卡尔曼滤波器。将非线性函数展开为泰勒多项式,仅保留线性项,这样积分又可以直接求解,得一种与LKF几乎相同的卡尔曼滤波器。但还要有一个额外的步骤,即需要计算每个非线性函数的雅可比矩阵。这种滤波器包括扩展卡尔曼滤波器(the Extended Kalman Filter,EKF)及其各种变形。

(3) Sigma点卡尔曼滤波器。将非线性函数展开为更一般的多项式,将积分简化为确定的矢量点集(称为Sigma点)的加权和。这类卡尔曼滤波器主要包括

有限差分卡尔曼滤波器(Finite Difference Kalman Filter,FDKF)、无迹卡尔曼滤波器(Unscented Kalman Filter,UKF)、球面单形卡尔曼滤波器(Spherical Simplex Kalman Filter,SSKF)和高斯厄密特卡尔曼滤波器(Gauss - Hermite Kalman Filter,GHKF)。

(4)蒙特卡洛卡尔曼滤波器。如果从高斯密度得一组蒙特卡洛采样点,生成离散密度,那么积分运算将简化为离散随机采样点求和运算。这种算法得蒙特卡洛卡尔曼滤波器(Monte Carlo Kalman Filter,MCKF)。

本书这一部分的很多材料是基于 Wu 等的论文[1]中的思想。Wu 等的论文简要介绍并比较了迄今为止提出的大多数的高斯加权积分解法,用一根主线把所有的卡尔曼滤波算法连接成一个统一的框架。从这篇论文当中,本书作者找到了很多数值积分方法相关文献,加深了对求解高斯加权积分的递归贝叶斯点估计方法的理解。本书这一部分的其他相关参考资料包括:McNamee 和 Stenger[2]的论文以及 Lemer[3]的博士论文,这二者的思想均源于 Wu 的文献;Davis 和 Rabinowitz[4]、Evans 和 Swartz[5]所著的数值积分书籍对于本书的完成起到重要作用;Ito 和 Xiong[6]的论文使作者对高斯滤波器基本概念有了初步了解,对完成本书具有重大意义。

5.1 第3章中重要结论总结

第3章的重要结论总结如下,将动态过程表示为

$$\boldsymbol{x}_n = \boldsymbol{f}_{n-1}(\boldsymbol{x}_{n-1}) + \boldsymbol{u}_n + \boldsymbol{v}_{n-1} \tag{5-1}$$

由观测过程生成的观测值为

$$z_n = \boldsymbol{h}_n(\boldsymbol{x}_n) + \boldsymbol{w}_n \tag{5-2}$$

基于包括 t_{n-1} 时刻在内的所有观测值,t_n 时刻状态矢量点预测估计由密度加权积分给出,即

$$\hat{\boldsymbol{x}}_{n|n-1} = \int_{\mathbf{R}^{n_x}} [\boldsymbol{f}_{n-1}(\boldsymbol{x}_{n-1}) + \boldsymbol{u}_n + \boldsymbol{v}_{n-1}] p(\boldsymbol{x}_{n-1} \mid z_{1:n-1}) \mathrm{d}\boldsymbol{x}_{n-1} \tag{5-3}$$

同样,基于包括 t_{n-1} 时刻在内的所有观测值,t_n 时刻的状态协方差的点预测估计由密度加权积分给出,即

$$\begin{aligned}\boldsymbol{P}^{xx}_{n|n-1} = \int_{\mathbf{R}^{n_x}} &[\boldsymbol{f}_{n-1}(\boldsymbol{x}_{n-1}) + \boldsymbol{u}_n - \boldsymbol{x}_{n|n-1}][\boldsymbol{f}_{n-1}(\boldsymbol{x}) + \boldsymbol{u}_n - \boldsymbol{x}_{n|n-1}]^{\mathrm{T}} \cdot \\ &p(\boldsymbol{x}_{n-1} \mid z_{1:n-1}) \mathrm{d}\boldsymbol{x}_{n-1} + \boldsymbol{Q}\end{aligned} \tag{5-4}$$

此外,观测矢量 z_n 的点预测估计、z_n 的协方差及其与 $\boldsymbol{x}_n$ 的互协方差由下式给出,即

$$\hat{z}_{n|n-1} = \int_{\mathbf{R}^{n_x}} [\boldsymbol{h}_n(\boldsymbol{x}_n) + \boldsymbol{w}_n] \, p(\boldsymbol{x}_n \mid z_{1:n-1}) \mathrm{d}\boldsymbol{x}_n \tag{5-5}$$

$$
\boldsymbol{P}_{n|n-1}^{zz} = \int_{\mathbf{R}^{n_x}} [\boldsymbol{h}_n(\boldsymbol{x}_n) - \hat{z}_{n|n-1}][\boldsymbol{h}_n(\boldsymbol{x}_n) - \hat{z}_{n|n-1}]^{\mathrm{T}}
$$

$$
p(\boldsymbol{x}_n \mid z_{1:n-1})\mathrm{d}\boldsymbol{x}_n + \boldsymbol{R} \tag{5-6}
$$

$$
\boldsymbol{P}_{n|n-1}^{xz} = \int_{\mathbf{R}^{n_x}} [\boldsymbol{x}_n - \hat{\boldsymbol{x}}_{n|n-1}][\boldsymbol{h}_n(\boldsymbol{x}_n) - \hat{z}_{n|n-1}]^{\mathrm{T}}
$$

$$
p(\boldsymbol{x}_n \mid z_{1:n-1})\mathrm{d}\boldsymbol{x}_n \tag{5-7}
$$

$$
\boldsymbol{Q} \triangleq \int_{\mathbf{R}^{n_x}} \boldsymbol{v}_{n-1}\boldsymbol{v}_{n-1}^{\mathrm{T}} p(\boldsymbol{x}_{n-1} \mid z_{1:n-1})\mathrm{d}\boldsymbol{x}_{n-1} \tag{5-8}
$$

$$
\boldsymbol{R} \triangleq \int_{\mathbf{R}^{n_x}}^{\mathbf{R}^{n_x}} \boldsymbol{w}_n\boldsymbol{w}_n^{\mathrm{T}} p(\boldsymbol{x}_n \mid z_{1:n-1})\mathrm{d}\boldsymbol{x}_n \tag{5-9}
$$

5.2 回顾卡尔曼滤波器校正(更新)方程的推导

基于所有条件密度均为高斯分布的假设，本节给出另一种卡尔曼滤波校正方程(3-43)、式(3-44)和式(3-47)的推导。贝叶斯法则提供的后验密度和 $\boldsymbol{x}_n$ 与 z_n 之间联合密度的关系为

$$
p(\boldsymbol{x}_n \mid \boldsymbol{z}_{1:n}) = \frac{p(\boldsymbol{x}_n, \boldsymbol{z}_n \mid z_{1:n-1})}{p(z_n)} \tag{5-10}
$$

定义联合矢量

$$
\boldsymbol{q} \triangleq \begin{bmatrix} \boldsymbol{x}_n \\ z_n \end{bmatrix} \tag{5-11}
$$

则式(5-10)可写为

$$
p(\boldsymbol{x}_n \mid z_{1:n}) = \frac{p(\boldsymbol{q}_n \mid z_{1:n-1})}{p(z_n)} \tag{5-12}
$$

多元高斯分布的一般形式可写为

$$
\mathcal{N}(\boldsymbol{t};\ \boldsymbol{s},\boldsymbol{\Sigma}) \triangleq \frac{1}{\sqrt{(2\pi)^n \mid \boldsymbol{\Sigma} \mid}}\exp\left\{-\frac{1}{2}(\boldsymbol{t}-\boldsymbol{s})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\boldsymbol{t}-\boldsymbol{s})\right\} \tag{5-13}
$$

令式(5-12)中的所有密度均为高斯分布，得

$$
p(\boldsymbol{x}_n \mid z_{1:n}) = \mathcal{N}(\boldsymbol{x}_n;\ \hat{\boldsymbol{x}}_{n|n}, \boldsymbol{P}_{n|n}^{xx}) \tag{5-14}
$$

$$
p(\boldsymbol{q}_n \mid z_{1:n-1}) = \mathcal{N}(\boldsymbol{q}_n;\ \hat{\boldsymbol{q}}_{n|n-1}, \boldsymbol{P}_{n|n-1}^{qq}) \tag{5-15}
$$

$$
p(z_n) = \mathcal{N}(z_n; \hat{z}_{n|n-1}, \boldsymbol{P}_{n|n-1}^{zz}) \tag{5-16}
$$

忽略(-1/2)因子，将式(5-14)中的后验高斯密度的指数解析式中的指数显式地写为

$$
\begin{aligned}
d &\triangleq (\boldsymbol{x}_n - \hat{\boldsymbol{x}}_{n|n})^{\mathrm{T}}[\boldsymbol{P}_{n|n}^{xx}]^{-1}(\boldsymbol{x}_n - \hat{\boldsymbol{x}}_{n|n}) \\
&= \boldsymbol{x}_n^{\mathrm{T}}[\boldsymbol{P}_{n|n}^{xx}]^{-1}\boldsymbol{x}_n - \boldsymbol{x}_n^{\mathrm{T}}[\boldsymbol{P}_{n|n}^{xx}]^{-1}\hat{\boldsymbol{x}}_{n|n} -
\end{aligned}
$$

$$x_{n|n}^{\mathrm{T}}[\boldsymbol{P}_{n|n}^{xx}]^{-1}x_n + x_{n|n}^{\mathrm{T}}[\boldsymbol{P}_{n|n}^{xx}]^{-1}\hat{x}_{n|n} \tag{5-17}$$

同样,式(5 - 15) 等号右边的高斯密度比率解析式中的指数为

$$\begin{aligned} g = &(q_n - \hat{q}_{n|n-1})^{\mathrm{T}}[\boldsymbol{P}_{n|n-1}^{qq}]^{-1}(q_n - \hat{q}_{n|n-1}) - \\ &(z_n - \hat{z}_{n|n-1})^{\mathrm{T}}[\boldsymbol{P}_{n|n-1}^{zz}]^{-1}(z_n - \hat{z}_{n|n-1}) \end{aligned} \tag{5-18}$$

由式(5 - 11),将式(5 - 18) 更显式地写为

$$\begin{aligned} g = &\begin{bmatrix} x_n - x_{n|n-1} \\ z_n - z_{n|n-1} \end{bmatrix}^{\mathrm{T}} \begin{bmatrix} \boldsymbol{P}_{n|n-1}^{xx} & \boldsymbol{P}_{n|n-1}^{xz} \\ \boldsymbol{P}_{n|n-1}^{zx} & \boldsymbol{P}_{n|n-1}^{zz} \end{bmatrix}^{-1} \begin{bmatrix} x_n - \hat{x}_{n|n-1} \\ z_n - \hat{z}_{n|n-1} \end{bmatrix} - \\ &(z_n - \hat{z}_{n|n-1})^{\mathrm{T}}[\boldsymbol{P}_{n|n-1}^{zz}]^{-1}(z_n - \hat{z}_{n|n-1}) \end{aligned} \tag{5-19}$$

定义式(5 - 19) 中分块矩阵项的逆为

$$\begin{bmatrix} \boldsymbol{C}_{11} & \boldsymbol{C}_{12} \\ \boldsymbol{C}_{21} & \boldsymbol{C}_{22} \end{bmatrix} \triangleq \begin{bmatrix} \boldsymbol{P}_{n|n-1}^{xx} & \boldsymbol{P}_{n|n-1}^{xz} \\ \boldsymbol{P}_{n|n-1}^{zx} & \boldsymbol{P}_{n|n-1}^{zz} \end{bmatrix}^{-1} \tag{5-20}$$

利用第 2. 1. 4 节[7] 中给出的分块矩阵的逆,得

$$\boldsymbol{C}_{11} = [\boldsymbol{P}_{n|n-1}^{xx} - \boldsymbol{P}_{n|n-1}^{xz}(\boldsymbol{P}_{n|n-1}^{zz})^{-1}\boldsymbol{P}_{n|n-1}^{zx}]^{-1} \tag{5-21}$$

$$\boldsymbol{C}_{12} = -\boldsymbol{C}_{11}\boldsymbol{P}_{n|n-1}^{xz}(\boldsymbol{P}_{n|n-1}^{zz})^{-1} \tag{5-22}$$

$$\boldsymbol{C}_{21} = -\boldsymbol{C}_{22}\boldsymbol{P}_{n|n-1}^{zx}(\boldsymbol{P}_{n|n-1}^{xx})^{-1} \tag{5-23}$$

$$\boldsymbol{C}_{22} = \boldsymbol{P}_{n|n-1}^{zz} - \boldsymbol{P}_{n|n-1}^{zx}(\boldsymbol{P}_{n|n-1}^{xx})^{-1}\boldsymbol{P}_{n|n-1}^{xz} \tag{5-24}$$

将式(5 - 20) ~ 式(5 - 24) 代入式(5 - 19),展开并整理,得

$$g = x_n^{\mathrm{T}}\boldsymbol{C}_{11}x_n + x_n^{\mathrm{T}}[-\boldsymbol{C}_{11}\hat{x}_{n|n-1} + \boldsymbol{C}_{12}(z_n - \hat{z}_{n|n-1})] + \cdots \tag{5-25}$$

将式(5 - 17) 的第 1 项与式(5 - 25) 第 1 项进行比较,得

$$\boldsymbol{P}_{n|n}^{xx} = \boldsymbol{P}_{n|n-1}^{xx} - \boldsymbol{P}_{n|n-1}^{xz}(\boldsymbol{P}_{n|n-1}^{zz})^{-1}\boldsymbol{P}_{n|n-1}^{zx} \tag{5-26}$$

将式(5 - 17) 的第 2 项与式(5 - 25) 第 2 项进行比较,得

$$\begin{aligned} [\boldsymbol{P}_{n|n}^{xx}]_{n|n}^{-1}\hat{x}_{n|n} &= \boldsymbol{C}_{11}\hat{x}_{n|n-1} - \boldsymbol{C}_{12}(z_n - z_{n|n-1}) \\ &= \boldsymbol{P}_{n|n-1}^{xx} - \boldsymbol{P}_{n|n-1}^{xz}(\boldsymbol{P}_{n|n-1}^{zz})^{-1}\boldsymbol{P}_{n|n-1}^{zx}]_{n|n-1}^{-1}\hat{x}_{n|n-1} + \\ &\quad [\boldsymbol{P}_{n|n-1}^{xx} - \boldsymbol{P}_{n|n-1}^{xz}(\boldsymbol{P}_{n|n-1}^{zz})^{-1}\boldsymbol{P}_{n|n-1}^{zx}]^{-1} \cdot \\ &\quad \boldsymbol{P}_{n|n-1}^{xz}(\boldsymbol{P}_{n|n-1}^{zz})^{-1}(z_n - \hat{z}_{n|n-1}) \end{aligned} \tag{5-27}$$

将式(5 - 27) 代入式(5 - 26),求解 $\hat{x}_{n|n}$,得

$$\hat{x}_{n|n} = \hat{x}_{n|n-1} + \boldsymbol{P}_{n|n-1}^{xz}(\boldsymbol{P}_{n|n-1}^{zz})^{-1}(z_n - \hat{z}_{n|n-1}) \tag{5-28}$$

$$\boldsymbol{K}_n \triangleq \boldsymbol{P}_{n|n-1}^{xz}(\boldsymbol{P}_{n|n-1}^{zz})^{-1} \tag{5-29}$$

现在,式(5 - 28) 和式(5 - 26) 变为卡尔曼滤波器校正方程,即

$$\hat{x}_{n|n} = \hat{x}_{n|n-1} + \boldsymbol{K}_n(z_n^{\mathrm{O}} - \hat{z}_{n|n-1}) \tag{5-30}$$

$$\boldsymbol{P}_{n|n}^{xx} = \boldsymbol{P}_{n|n-1}^{xx} - \boldsymbol{K}_n\boldsymbol{P}_{n|n-1}^{zz}\boldsymbol{K}_n^{\mathrm{T}} \tag{5-31}$$

这里用实际的观测值 z_n^{O} 代替 z_n,并且用到 $\boldsymbol{P}_{n|n-1}^{xz}(\boldsymbol{P}_{n|n-1}^{zz})^{-1}\boldsymbol{P}_{n|n-1}^{zx} = \boldsymbol{K}_n\boldsymbol{P}_{n|n-1}^{zz}\boldsymbol{K}_n^{\mathrm{T}}$ 这一事实。注意在推导卡尔曼滤波器校正方程过程中,对式(5 - 1) 动态方程和式(5 - 2) 观测方程的线性度没有做任何假设。

因此，卡尔曼滤波校正方程可由高斯分布直接推导得，而无需任何假设或条件。然而，记住这一点很重要，因为这些校正方程比较通用，可以应用于任意分布的前两阶矩（只要这些矩存在）。必须指出，到目前为止没有期刊文献将这些校正方程用在非高斯密度当中。

5.3 贝叶斯点预测积分求解高斯密度

根据动态方程（5－1），假定 $\boldsymbol{v}_n$ 为独立于 $\boldsymbol{x}_n \in \mathbf{R}^{n_x}$ 和 $\boldsymbol{z}_n \in \mathbf{R}^{n_z}$ 的零均值高斯随机变量，得

$$\boldsymbol{v}_{n-1} \sim p(\boldsymbol{x}_n \mid \boldsymbol{x}_{n-1}) = \mathcal{N}(\boldsymbol{x}_n - [\boldsymbol{f}_{n-1}(\boldsymbol{x}_{n-1}) + \boldsymbol{u}_n]; \boldsymbol{0}, \boldsymbol{Q}) = \mathcal{N}(\boldsymbol{v}_{n-1}; \boldsymbol{0}, \boldsymbol{Q}) \tag{5-32}$$

同样，对于观测方程（5－2），假定 $\boldsymbol{w}_n$ 是一个独立于 $\boldsymbol{x}_n$ 和 $\boldsymbol{z}_n$ 的零均值高斯随机变量，有

$$\boldsymbol{w}_n \sim p(\boldsymbol{z}_n \mid \boldsymbol{x}_n) = \mathcal{N}(\boldsymbol{z}_n - \boldsymbol{h}_n(\boldsymbol{x}_n); \boldsymbol{0}, \boldsymbol{R}) = \mathcal{N}(\boldsymbol{w}_n; \boldsymbol{0}, \boldsymbol{R}) \tag{5-33}$$

由贝叶斯法则，得先验密度，即

$$p(\boldsymbol{x}_n \mid \boldsymbol{z}_{1:n-1}) \propto p(\boldsymbol{z}_n \mid \boldsymbol{x}_n) p(\boldsymbol{x}_n \mid \boldsymbol{x}_{n-1}) \tag{5-34}$$

这里假设所有的密度函数均为高斯分布，令 $p(\boldsymbol{x}_n \mid \boldsymbol{z}_{1:n-1})$ 为高斯分布，因为两个高斯密度的积同样是高斯分布，则有

$$p(\boldsymbol{x}_n \mid \boldsymbol{z}_{1:n-1}) = \mathcal{N}(\boldsymbol{x}_n; \hat{\boldsymbol{x}}_{n|n-1}, \boldsymbol{P}^{xx}_{n|n-1}) \tag{5-35}$$

$$p(\boldsymbol{x}_{n-1} \mid \boldsymbol{z}_{1:n-1}) = \mathcal{N}(\boldsymbol{x}_{n-1}; \hat{\boldsymbol{x}}_{n-1|n-1}, \boldsymbol{P}^{xx}_{n-1|n-1}) \tag{5-36}$$

假设 $\boldsymbol{v}_n$ 和 $\boldsymbol{w}_n$ 为零均值高斯过程，将式（5－36）代入式（5－3）和式（5－4），式（5－35）代入式（5－5）～式（5－7），预测估计过程可以用高斯加权积分来概括，即

$$\hat{\boldsymbol{x}}_{n|n-1} = \int \boldsymbol{f}_{n-1}(\boldsymbol{x}_{n-1}) \cdot \mathcal{N}(\boldsymbol{x}_{n-1}; \hat{\boldsymbol{x}}_{n-1|n-1}, \boldsymbol{P}^{xx}_{n-1|n-1}) \mathrm{d}\boldsymbol{x}_{n-1} + \boldsymbol{u}_n \tag{5-37}$$

$$\begin{aligned}\boldsymbol{P}^{xx}_{n|n-1} = \int & [\boldsymbol{f}_{n-1}(\boldsymbol{x}_{n-1}) + \boldsymbol{u}_n - \hat{\boldsymbol{x}}_{n|n-1}] \cdot \\ & [\boldsymbol{f}_{n-1}(\boldsymbol{x}_{n-1}) + \boldsymbol{u}_n - \hat{\boldsymbol{x}}_{n|n-1}]^{\mathrm{T}} \cdot \\ & \mathcal{N}(\boldsymbol{x}_{n-1}; \hat{\boldsymbol{x}}_{n-1|n-1}, \boldsymbol{P}^{xx}_{n-1|n-1}) \mathrm{d}\boldsymbol{x}_{n-1} + \boldsymbol{Q}\end{aligned} \tag{5-38}$$

$$\hat{\boldsymbol{z}}_{n|n-1} = \int \boldsymbol{h}_n(\boldsymbol{x}_n) \mathcal{N}(\boldsymbol{x}_n; \hat{\boldsymbol{x}}_{n|n-1}, \boldsymbol{P}^{xx}_{n|n-1}) \mathrm{d}\boldsymbol{x}_n \tag{5-39}$$

$$\begin{aligned}\boldsymbol{P}^{zz}_{n|n-1} = \int & (\boldsymbol{h}_n(\boldsymbol{x}_n) - \hat{\boldsymbol{z}}_{n|n-1})(\boldsymbol{h}_n(\boldsymbol{x}_n) - \hat{\boldsymbol{z}}_{n|n-1})^{\mathrm{T}} \cdot \\ & \mathcal{N}(\boldsymbol{x}_n; \hat{\boldsymbol{x}}_{n|n-1}, \boldsymbol{P}^{xx}_{n|n-1}) \mathrm{d}\boldsymbol{x}_n + \boldsymbol{R}\end{aligned} \tag{5-40}$$

$$P_{n|n-1}^{xz} = \int (\boldsymbol{x}_n - \hat{\boldsymbol{x}}_{n|n-1})(\boldsymbol{h}_n(\boldsymbol{x}_n) - \hat{\boldsymbol{z}}_{n|n-1})^{\mathrm{T}} \cdot \mathcal{N}(\boldsymbol{x}_n; \hat{\boldsymbol{x}}_{n|n-1}, \boldsymbol{P}_{n|n-1}^{xx}) \mathrm{d}\boldsymbol{x}_n \tag{5-41}$$

$$\boldsymbol{R} \triangleq \int_{\mathbf{R}^{n_x}} \boldsymbol{w}_n \boldsymbol{w}_n^{\mathrm{T}} \mathcal{N}(\boldsymbol{w}_n; \boldsymbol{0}, \boldsymbol{R}) \mathrm{d}\boldsymbol{w}_n \tag{5-42}$$

$$\boldsymbol{Q} \triangleq \int_{\mathbf{R}^{n_x}} \boldsymbol{v}_{n-1} \boldsymbol{v}_{n-1}^{\mathrm{T}} \mathcal{N}(\boldsymbol{v}_{n-1}; \boldsymbol{0}, \boldsymbol{Q}) \mathrm{d}\boldsymbol{v}_{n-1} \tag{5-43}$$

图5－1给出了高斯概率分布下一般贝叶斯估计的流程框图，包括以下步骤：

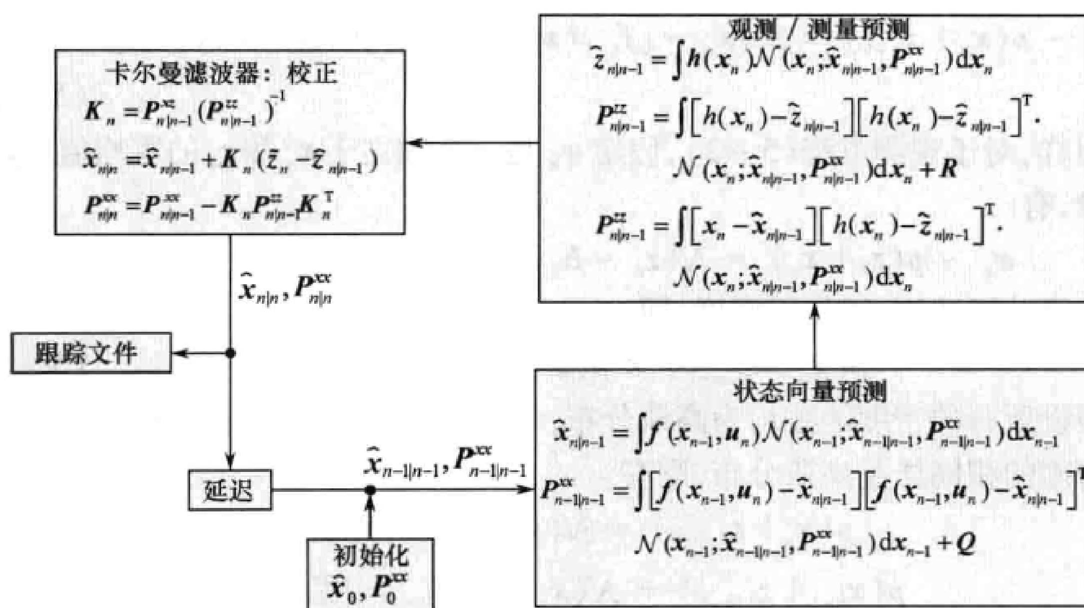


图5－1　高斯概率密度贝叶斯估计流程框图

（1）根据初步观测，生成 $\hat{\boldsymbol{x}}_0$ 和 $\boldsymbol{P}_0^{xx}$ 的最佳猜测值，初始化跟踪滤波器；

（2）计算状态矢量预测积分；

（3）计算观测值预测积分；

（4）执行卡尔曼滤波器校正方程，得后验点估计 $\hat{\boldsymbol{x}}_{n|n}$ 和 $\boldsymbol{P}_{n|n}^{xx}$；

（5）将结果存储在跟踪文件中；

（6）时间上前进一步，利用 $\hat{\boldsymbol{x}}_{n-1|n-1}$ 和 $\boldsymbol{P}_{n-1|n-1}^{xx}$ 作为新的输入，进入下一轮滤波器迭代；

（7）返回步骤（2）。

所有基于高斯假设的贝叶斯估计方法都遵循同样的基本流程，只是在函数逼近或者积分方法上有所改变。

5.3.1　利用仿射变换来简化流程

第2.4.2.3节讨论了将矢量仿射变换应用到式（2－106）形式的一般高斯积

分当中。清晰起见,这里再重复一次,即

$$\varepsilon\{f(\boldsymbol{x})\} = \int f(\boldsymbol{x}) \frac{1}{(2\pi)^{n_x/2} \mid \boldsymbol{P}^{xx} \mid^{1/2}} \cdot \exp\left\{-\frac{1}{2}(\boldsymbol{x}-\hat{\boldsymbol{x}})^{\mathrm{T}}[\boldsymbol{P}^{xx}]^{-1}(\boldsymbol{x}-\hat{\boldsymbol{x}})\right\}\mathrm{d}\boldsymbol{x} \tag{5-44}$$

$$= \int f(\boldsymbol{x})\mathcal{N}(\boldsymbol{x};\hat{\boldsymbol{x}},\boldsymbol{P}^{xx})\,\mathrm{d}\boldsymbol{x} \tag{5-45}$$

如第 2 章内容所述进行矢量仿射变换

$$\boldsymbol{x} = \hat{\boldsymbol{x}} + \boldsymbol{D}\boldsymbol{c} \tag{5-46}$$

则得

$$\int f(\boldsymbol{x})\mathcal{N}(\boldsymbol{x};\hat{\boldsymbol{x}},\boldsymbol{P}^{xx})\,\mathrm{d}\boldsymbol{x} = \int \tilde{\boldsymbol{f}}(\boldsymbol{c})\mathcal{N}(\boldsymbol{c};\boldsymbol{0},\boldsymbol{I})\,\mathrm{d}\boldsymbol{c} \tag{5-47}$$

$\boldsymbol{D}$ 由下式定义

$$\boldsymbol{P}^{xx} = \boldsymbol{D}\boldsymbol{D}^{\mathrm{T}} \tag{5-48}$$

或者,可写为

$$\boldsymbol{D} = [\boldsymbol{P}^{xx}]^{1/2} \tag{5-49}$$

此外,$\tilde{f}(\boldsymbol{c})$ 将定义为

$$\tilde{\boldsymbol{f}}(\boldsymbol{c}) \triangleq \boldsymbol{f}(\hat{\boldsymbol{x}} + \boldsymbol{D}\boldsymbol{c}) \tag{5-50}$$

对式(5-37)、式(5-38)应用矢量仿射变换,得

$$\hat{\boldsymbol{x}}_{n|n-1} = \int \tilde{\boldsymbol{f}}(\boldsymbol{c})\mathcal{N}(\boldsymbol{c};\boldsymbol{0},\boldsymbol{I})\,\mathrm{d}\boldsymbol{c} + \boldsymbol{u}_n \tag{5-51}$$

$$\boldsymbol{P}^{xx}_{n|n-1} = \int [\tilde{\boldsymbol{f}}(\boldsymbol{c}) + \boldsymbol{u}_n - \hat{\boldsymbol{x}}_{n|n-1}][\tilde{\boldsymbol{f}}(\boldsymbol{c}) + \boldsymbol{u}_n - \hat{\boldsymbol{x}}_{n|n-1}]^{\mathrm{T}} \cdot \mathcal{N}(\boldsymbol{c};\boldsymbol{0},\boldsymbol{I})\,\mathrm{d}\boldsymbol{c} + \boldsymbol{Q} \tag{5-52}$$

$$\tilde{\boldsymbol{f}}(\boldsymbol{c}) = \tilde{\boldsymbol{f}}(\hat{\boldsymbol{x}}_{n-1|n-1} + \boldsymbol{D}_{n-1|n-1}\boldsymbol{c}) \tag{5-53}$$

$$\boldsymbol{D}_{n-1|n-1} \triangleq [\boldsymbol{P}^{xx}_{n-1|n-1}]^{1/2} \tag{5-54}$$

此外,式(5-39)~式(5-41)变成

$$\hat{\boldsymbol{z}}_{n|n-1} = \int \tilde{\boldsymbol{h}}(\boldsymbol{c})\mathcal{N}(\boldsymbol{c};\boldsymbol{0},\boldsymbol{I})\,\mathrm{d}\boldsymbol{c} \tag{5-55}$$

$$\boldsymbol{P}^{zz}_{n|n-1} = \int [\tilde{\boldsymbol{h}}(\boldsymbol{c}) - \hat{\boldsymbol{z}}_{n|n-1}][\tilde{\boldsymbol{h}}(\boldsymbol{c}) - \hat{\boldsymbol{z}}_{n|n-1}]^{\mathrm{T}} \cdot \mathcal{N}(\boldsymbol{c};\boldsymbol{0},\boldsymbol{I})\,\mathrm{d}\boldsymbol{c} + \boldsymbol{R} \tag{5-56}$$

$$= \int [\tilde{\boldsymbol{f}}(\boldsymbol{c}) + \boldsymbol{x}_n - \hat{\boldsymbol{x}}_{n|n-1}][\hat{\boldsymbol{h}}(\boldsymbol{c}) - \hat{\boldsymbol{z}}_{n|n-1}]\mathcal{N}(\boldsymbol{c};\boldsymbol{0},\boldsymbol{I})\,\mathrm{d}\boldsymbol{c} \tag{5-57}$$

$$\tilde{\boldsymbol{h}}(\boldsymbol{c}) = \boldsymbol{h}(\hat{\boldsymbol{x}}_{n|n-1} + \boldsymbol{D}_{n|n-1}\boldsymbol{c}) \tag{5-58}$$

$$\boldsymbol{D}_{n|n-1} = [\boldsymbol{P}^{xx}_{n|n-1}]^{1/2} \tag{5-59}$$

式(5 - 51) ~ 式(5 - 52) 和式(5 - 55) ~ 式(5 - 57) 方程组需要求解三个如下形式的一般矩积分,即

$$\boldsymbol{I}(\tilde{\boldsymbol{f}}) \triangleq \int \tilde{\boldsymbol{f}}(\boldsymbol{c}) \mathcal{N}(\boldsymbol{c};\boldsymbol{0},\boldsymbol{I}) \mathrm{d}\boldsymbol{c} \tag{5-60}$$

$$\boldsymbol{I}(\tilde{\boldsymbol{f}}\tilde{\boldsymbol{f}}^{\mathrm{T}}) \triangleq \int [\tilde{\boldsymbol{f}}(\boldsymbol{c}) - \boldsymbol{I}(\tilde{\boldsymbol{f}})][\tilde{\boldsymbol{f}}(\boldsymbol{c}) - \boldsymbol{I}(\tilde{\boldsymbol{f}})]^{\mathrm{T}} \mathcal{N}(\boldsymbol{c};\boldsymbol{0},\boldsymbol{I}) \mathrm{d}\boldsymbol{c} \tag{5-61}$$

$$\boldsymbol{I}(\boldsymbol{c}\tilde{\boldsymbol{f}}) \triangleq \int \boldsymbol{c}\tilde{\boldsymbol{f}}(\boldsymbol{c}) \mathcal{N}(\boldsymbol{c};\boldsymbol{0},\boldsymbol{I}) \mathrm{d}\boldsymbol{c} \tag{5-62}$$

5.3.2 求解高斯加权积分的一般方法

对于一般非线性函数 $\tilde{\boldsymbol{f}}(\boldsymbol{c})$,无法求出式(5 - 60) ~ 式(5 - 62) 中积分的解析解。一般将求解这些积分的数值方法称为多重积分规则,每种规则可以求解一类 $\tilde{\boldsymbol{f}}(\boldsymbol{c})$ 多维多项式逼近的积分。在文献[8] 及其引用的参考文献中,均采用了类似方法。这种方法通过对随机变量 $\boldsymbol{x}$ 的一系列变换,对 $\boldsymbol{f}(\boldsymbol{x})$ 进行统计线性化处理,然后将变换后的 $\boldsymbol{f}(\boldsymbol{x})$ 进行多项式展开,计算期望积分。

式(5 - 46) 的仿射变换非常重要,因为它将一般高斯加权积分变换为完全对称。密度加权积分完全对称就是在均等轮廓形成的超球面所有维度上,密度关于零对称。此外,在所有维度上,积分的上下限也必须关于零对称。这将大大简化用高斯加权积分对非线性函数进行近似的数值方法。

根据解析式 $\mathcal{N}(\boldsymbol{c};\boldsymbol{0},\boldsymbol{I})$ 可以看出,对于多维坐标系统下的矢量点 $\boldsymbol{c}^{(i)}$,$\mathcal{N}(\boldsymbol{c}^{(i)};\boldsymbol{0},\boldsymbol{I})$ 对应具体值。对于 n_x 维单位半径超球面上任意 n_x 维矢量点 $\boldsymbol{c}^{(i)}$,可直接得出 $\mathcal{N}(\boldsymbol{c}^{(i)};\boldsymbol{0},\boldsymbol{I}) = \mathrm{e}^{-1/2}/(2\pi)^{-n_x/2}$。以上还可以推广为:对于一个给定半径的超球面,对于超球面表面上的任意矢量点 $\boldsymbol{c}^{(i)}$,$\mathcal{N}(\boldsymbol{c}^{(i)};\boldsymbol{0},\boldsymbol{I})$ 取值相同。在二维的情况下,可以由图 2 - 9 中看到恒定半径 $\mathcal{N}(\boldsymbol{c};0,\boldsymbol{I})$ 的等值轮廓圆。

假定 $\tilde{\boldsymbol{f}}(\boldsymbol{c})$ 是线性的,即 $\tilde{\boldsymbol{f}}(\boldsymbol{c}) \to \boldsymbol{F}\boldsymbol{c}$,其中 $\boldsymbol{F}$ 为独立于 $\boldsymbol{c}$ 的确定性转移矩阵,那么式(5 - 60) ~ 式(5 - 62) 的求解会非常简单,式(5 - 51) ~ 式(5 - 57) 将简化为著名的线性卡尔曼滤波器,在第 6 章会详细介绍。

如果 $\tilde{\boldsymbol{f}}(\boldsymbol{c})$ 为非线性,常规的方法是将 $\tilde{\boldsymbol{f}}(\boldsymbol{c})$ 解析展开为一个多维多项式,每一维都由单项式的乘积构成。例如,对于 $\boldsymbol{c} = [c_1, c_2]^{\mathrm{T}}$,$\tilde{\boldsymbol{f}}(\boldsymbol{c})$ 的二维多项式展开为

$$
\begin{aligned}
\tilde{\boldsymbol{f}}(c_1,c_2) &= \sum_{i_1=0}^{m_1}\sum_{i_2=0}^{m_2} a(i_1,i_2)c_1^{i_1}c_2^{i_2} \\
&= a(0,0) + a(1,0)c_1 + a(0,1)c_2 + a(1,1)c_1c_2 + a(2,0)c_1^2 \\
&\quad + \cdots + a(m_1,m_2)c_1^{m_1}c_2^{m_2}
\end{aligned} \tag{5-63}
$$

多项式的每一项都是 $a(i_1,i_2)c_1^{i_1}c_2^{i_2}$ 形式的单项式。对于 d 次多维多项式，每一项都包含单项式 $\prod_{j=1}^{d} c_j^{i_j}$，其中 $i_j \geqslant 0$ 且 $\sum_{j=1}^{d} i_j \leqslant d$。对于上述二维情况，多项式的次数为 $d = m_1 + m_2 \geqslant \sum_{j=1}^{d} i_j$。

对于非线性函数 $\tilde{\boldsymbol{f}}(\boldsymbol{c})$，用 $\tilde{\boldsymbol{f}}(\boldsymbol{c})$ 的积分多项式展开为形如式(5-60)～式(5-62)的方程，得一般多重积分规则。下面介绍积分规则[1,9]的 p 精度，对于式(5-60)、式(5-62)形式的积分，"如果一个规则的积分单项式最高为 p 次，那么就说它具有 p 精度，即：单项式形如 $\prod_{j=1}^{p} c_j^{i_j}$，$i_j \geqslant 0$ 且 $\sum_{j=1}^{p} i_j \leqslant p$，但对一些次数为 $\sum_{j=1}^{p} i_j \leqslant p+1$ 的单项式并不适用"[1]。

多项式展开的经典方法是将 $\tilde{\boldsymbol{f}}(\boldsymbol{c})$ 展开为泰勒多项式。在第 7 章，仅保留泰勒多项式的线性项，得扩展卡尔曼滤波器（EKF）。然而，EKF 的主要缺陷是需要求解雅可比矩阵。对于高非线性度函数，使用包含二阶项的泰勒多项式可以得更准确的解，但它需要额外求解 Hessian 矩阵。这种求解方法可推导出二阶 EKF，在第 7 章中会详细介绍。

另一种替代方法是将雅可比和 Hessian 差分矩阵替换为多维中心有限差分近似，导出 FDKF。FDKF 是第 8 章的主题内容，是 Sigma 点卡尔曼滤波器的第一个例子。Sigma 点卡尔曼滤波器需要求解在构成多维网格的 Sigma 点上 $\tilde{\boldsymbol{f}}(\boldsymbol{c})$ 的值。此外，对于 Sigma 点卡尔曼滤波器，求解方程式(5-60)～式(5-62)均要解一组矩方程，在后续章节详细讨论。在数值积分方法的文献中，这个概念方法被推广到将式(5-60)近似为如下形式[10]，即

$$
\boldsymbol{I}(\tilde{\boldsymbol{f}}) = \sum_{j=1}^{n_s} w_j \tilde{\boldsymbol{f}}(\boldsymbol{c}^{(j)}), \tilde{\boldsymbol{f}}(\boldsymbol{c}) \in \mathbf{R}^{n_x} \tag{5-64}
$$

式中：权值 ω_j、Sigma 点 $\boldsymbol{c}^{(j)}$ 和点的数目 n_s 由所选取的积分方法决定。ω_j 和 $\boldsymbol{c}^{(j)}$ 的取值与函数 $\tilde{\boldsymbol{f}}$ 无关。如果积分域的维数为 1，即 $n_x = 1$，这种近似称为求积公式。如果 $n_x \geqslant 2$，这种近似叫做求体积公式具体可分为：数论方法和基于多项式的方法。第一类要求积分（Sigma）点在多维网格上均匀分布；第二类方法对于一些正交多项式集合非常精确，FDKF 属于第一类。

下面要使用一个超球面上的两组矢量点。对于第一组，假定矢量点 $\boldsymbol{c}^{(i)} =$

$q[1,0,\cdots,0]^{\mathrm{T}} = q\boldsymbol{r}^{(i)}$，其中 q 为超球面的半径、$\boldsymbol{r}^{(i)}$ 为沿某一坐标轴的单位矢量。超球面上其它沿不同坐标轴的点可以通过对第一个点改变符号和／或交换维度来得。例如，点 $\boldsymbol{c}^{(j)} = q[-1,0,\cdots,0]^{\mathrm{T}}$ 或 $\boldsymbol{c}^{(j)} = q[0,\cdots,1,\cdots,0]^{\mathrm{T}}$ 都和初始点一样，在同一个超球面上。实际上，在每根坐标轴上有两个点，一个为正、一个为负，这样一共有 $2n_x$ 个类似矢量点。将这个矢量点集合定义为 $q\boldsymbol{r}$，$\boldsymbol{r}$ 为沿 n_x 笛卡儿坐标轴的正、负方向的单位矢量。这种方法得出了 UKF，将在第 9 章讨论。

对称性也可以用超球面上等角间距（而不是沿着坐标轴）的矢量点集来表示。所有这些在一个超球面上矢量点将形成一个对称集合，从原点辐射到多维单形的顶点。如果给定一个点以及点之间的旋转角，那么可以由初始点得集合中的所有点。可以看到，这个集合仅包含 n_x+2 个点。由这个方法得出 SSKF，在第 10 章会对此详加阐述。

实际上有许多这样的点集可以作为生成器。比如，在一个方形网格上（而不是超球面上）的两个矢量点 $\boldsymbol{c}^{(i)} = q[1,1,0,\cdots,0]^{\mathrm{T}}$ 和 $\boldsymbol{c}^{(i)} = q[1,1,1,\cdots,1]^{\mathrm{T}}$，每个都可以作为一个生成器，生成一组不同的对称点集。但是在这些情况下，所有的 $\mathcal{N}(\boldsymbol{c}^{(i)};\boldsymbol{0},\boldsymbol{I})$ 值并不一定都相等。这种矢量点生成器将在第 11 章中介绍，还将用正交厄密特多项式得出 GHKF。

由于对多项式次数有限制，对 $\tilde{\boldsymbol{f}}(\boldsymbol{c})$ 进行多项式逼近，必须满足如下矩方程组

$$\sum_{j=1}^{n_s} w_j \tilde{\boldsymbol{f}}_i(\boldsymbol{c}^{(j)}) = \int \tilde{\boldsymbol{f}}_i(\boldsymbol{c})\mathcal{N}(\boldsymbol{c};\boldsymbol{0},\boldsymbol{I})\mathrm{d}\boldsymbol{c} \tag{5-65}$$

式中：$\tilde{\boldsymbol{f}}_i(\boldsymbol{c})$ 为对 $\tilde{\boldsymbol{f}}(\boldsymbol{c})$ 进行多项式展开后，次数为 i 的项。由矩方程组得的非线性方程组可以用来求解未知项 $\{\omega_j,\boldsymbol{c}^{(j)}\}$。式(5－61) 的矩阵方程变成

$$\sum_{j=1}^{n_s} w_j \tilde{\boldsymbol{f}}_i(\boldsymbol{c}^{(j)})\tilde{\boldsymbol{f}}_l(\boldsymbol{c}^{(j)}) = \int \tilde{\boldsymbol{f}}_i(\boldsymbol{c})\tilde{\boldsymbol{f}}_l(\boldsymbol{c})\mathcal{N}(\boldsymbol{c};\boldsymbol{0},\boldsymbol{I})\mathrm{d}\boldsymbol{c} \tag{5-66}$$

如果想进一步详细了解这些数值方法，可参考 Davis 和 Rabinowitz[4]、Evans 和 Swartz[5] 的著作以及其它参考材料。

参考文献

[1] Wu Y, Hu D, Wu M, Hu X. A numerical－integration perspective on Gaussian filters. *IEEE Trans. Sig. Proc*, 2006, 54(8): 2910－2921.

[2] McNamee J, Stenger F. Construction of fully symmetric numerical integration formulas. Numerische Math, 1967, 10: 327－344.

[3] Lerner UN. Hybrid Bayesian Networks for Reasoning About Complex Systems, Dissertation. Department

of Computer Science, Stanford University, October 2002.

[4] Davis PJ, Rabinowitz P. *Methods of Numerical Integration.* Reprint of Second Edition. Minneola, NY: Dover Publications, 1984.

[5] Evans M, Swartz T. *Approximating Integrals via Monte Carlo and Deterministic Methods.* Oxford Statistical Science Series: 20 Reprinted 2005. Oxford University Press, 2000.

[6] Ito K, Xiong K. Gaussian filters for nonlinear filtering problems. *IEEE Trans. Automatic Control*, 2000, 45(5):910 - 927.

[7] Marple SL. Digital Spectral Analysis with Applications. Prentice - Hall, 1987.

[8] Gelb A. *Applied Optimal Estimation.* The MIT Press, 1974.

[9] Stroud AK. *Approximate Calculation of Multiple Integrals.* Englewood cliffs, NJ: Prentice Hall, 1971.

[10] Cool R. Advances in multidimensional integration. *J. Comput. and Appl. Math*, 2002, 149:1 - 12.

第6章 线性卡尔曼滤波器

递归线性卡尔曼滤波器(LKF)由卡尔曼[1,2]提出,在文献[3-5]中结合应用实例进行了深入讨论,这些书的新版本或原书的再版现在都可以买到。很多估计问题动态方程或观测方程是线性的,有时二者均为线性。因此,这里将单独处理动态方程或观测方程为线性的情况,将其归为线性卡尔曼滤波器。

6.1 线性动态模型

对于线性动态模型,式(5-1)可写成

$$\boldsymbol{x}_n = \boldsymbol{F}\boldsymbol{x}_{n-1} + \boldsymbol{u}_n + \boldsymbol{v}_{n-1} \tag{6-1}$$

式中:$\boldsymbol{F}$ 为独立于时间的确定性转移矩阵。多数情况下,不需要控制因子 $\boldsymbol{u}_n$,但为确保完整,后面将包含这部分内容。

将式(6-1)代入式(5-37),得

$$\hat{\boldsymbol{x}}_{n|n-1} = \boldsymbol{F}\int \boldsymbol{x}_{n-1}\mathcal{N}(\boldsymbol{x}_{n-1};\hat{\boldsymbol{x}}_{n-1|n-1},\boldsymbol{P}^{xx}_{n-1|n-1})\mathrm{d}\boldsymbol{x}_{n-1} + \boldsymbol{u}_n \tag{6-2}$$

进一步简化为

$$\hat{\boldsymbol{x}}_{n|n-1} = \boldsymbol{F}\hat{\boldsymbol{x}}_{n-1|n-1} + \boldsymbol{u}_n \tag{6-3}$$

同样,将式(6-1)代入式(5-38),得

$$\begin{aligned}\boldsymbol{P}^{xx}_{n|n-1} &= \int \boldsymbol{F}[\boldsymbol{x}_{n-1} - \hat{\boldsymbol{x}}_{n|n-1}][\boldsymbol{x}_{n-1} - \hat{\boldsymbol{x}}_{n|n-1}]^{\mathrm{T}}\boldsymbol{F}^{\mathrm{T}} \cdot \\ &\mathcal{N}(\boldsymbol{x}_{n-1};\hat{\boldsymbol{x}}_{n-1|n-1},\boldsymbol{P}^{xx}_{n-1|n-1})\mathrm{d}\boldsymbol{x}_{n-1} + \boldsymbol{Q} \\ &= \boldsymbol{F}\left\{\int [\boldsymbol{x}_{n-1} - \hat{\boldsymbol{x}}_{n|n-1}][\boldsymbol{x}_{n-1} - \hat{\boldsymbol{x}}_{n|n-1}]^{\mathrm{T}} \cdot \right. \\ &\left.\mathcal{N}(\boldsymbol{x}_{n-1};\hat{\boldsymbol{x}}_{n-1|n-1},\boldsymbol{P}^{xx}_{n-1|n-1})\mathrm{d}\boldsymbol{x}_{n-1}\right\}\boldsymbol{F}^{\mathrm{T}} + \boldsymbol{Q}\end{aligned} \tag{6-4}$$

但是,由于

$$\boldsymbol{P}^{xx}_{n-1|n-1} = \int [\boldsymbol{x}_{n-1} - \hat{\boldsymbol{x}}_{n|n-1}][\boldsymbol{x}_{n-1} - \hat{\boldsymbol{x}}_{n|n-1}]^{\mathrm{T}} \cdot$$

$$\mathcal{N}(\boldsymbol{x}_{n-1};\hat{\boldsymbol{x}}_{n-1|n-1},\boldsymbol{P}^{xx}_{n-1|n-1})\mathrm{d}\boldsymbol{x}_{n-1} \tag{6-5}$$

对于线性动态模型,有

$$\boldsymbol{P}^{xx}_{n|n-1} = \boldsymbol{F}\boldsymbol{P}^{xx}_{n-1|n-1}\boldsymbol{F}^{\mathrm{T}} + \boldsymbol{Q}$$
$$\boldsymbol{Q} = \varepsilon(\boldsymbol{v}_{n-1}\boldsymbol{v}^{\mathrm{T}}_{n-1}) \tag{6-6}$$

式中:$\boldsymbol{Q}$ 为动态噪声协方差矩阵。

6.2 线性观测模型

如果观测方程为线性,可将观测方程写成

$$\boldsymbol{z}_n = \boldsymbol{H}\boldsymbol{x}_n + \boldsymbol{w}_n \tag{6-7}$$

式中:$\boldsymbol{H}$ 为 $n_z \times n_x$ 维,且与 $\boldsymbol{x}_n$ 和时间无关的确定性转移矩阵。因此有

$$\boldsymbol{h}(\boldsymbol{x}_n) = \boldsymbol{H}\boldsymbol{x}_n \tag{6-8}$$

假定观测噪声 $\boldsymbol{w}_n$ 为零均值过程,则式(5-39)和式(5-40)可简化为

$$\hat{\boldsymbol{z}}_{n|n-1} = \boldsymbol{H}\hat{\boldsymbol{x}}_{n|n-1} \tag{6-9}$$

$$\boldsymbol{P}^{zz}_{n|n-1} = \boldsymbol{H}\boldsymbol{P}^{xx}_{n|n-1}\boldsymbol{H}^{\mathrm{T}} + \boldsymbol{R}$$
$$\boldsymbol{R} = \varepsilon(\boldsymbol{w}_n\boldsymbol{w}^{\mathrm{T}}_n) \tag{6-10}$$

式中:R 为观测噪声协方差矩阵。

最后,将式(6-7)、式(6-9)代入式(5-57),得

$$\begin{aligned}
\boldsymbol{P}^{xz}_{n|n-1} &= \int[\boldsymbol{x}_n-\hat{\boldsymbol{x}}_{n|n-1}][\boldsymbol{z}_n-\hat{\boldsymbol{z}}_{n|n-1}]^{\mathrm{T}}\mathcal{N}(\boldsymbol{x}_n;\hat{\boldsymbol{x}}_{n|n-1},\boldsymbol{P}^{xx}_{n|n-1})\mathrm{d}\boldsymbol{x}_n \\
&= \int[\boldsymbol{x}_n-\boldsymbol{x}_{n|n-1}][\boldsymbol{H}\boldsymbol{x}_n+\boldsymbol{w}_n-\boldsymbol{H}\hat{\boldsymbol{x}}_{n|n-1}]^{\mathrm{T}}\cdot \\
&\quad \mathcal{N}(\boldsymbol{x}_n;\hat{\boldsymbol{x}}_{n|n-1},\boldsymbol{P}^{xx}_{n|n-1})\mathrm{d}\boldsymbol{x}_n \\
&= \left\{\int[\boldsymbol{x}_n-\hat{\boldsymbol{x}}_{n|n-1}][\boldsymbol{x}_n-\hat{\boldsymbol{x}}_{n|n-1}]^{\mathrm{T}}\right. \\
&\quad \left.\mathcal{N}(\boldsymbol{x}_n;\hat{\boldsymbol{x}}_{n|n-1},\boldsymbol{P}^{xx}_{n|n-1})\mathrm{d}\boldsymbol{x}_n\right\}\boldsymbol{H}^{\mathrm{T}} + \\
&\quad \int[\boldsymbol{x}_n-\hat{\boldsymbol{x}}_{n|n-1}]\boldsymbol{w}^{\mathrm{T}}_n\mathcal{N}(\boldsymbol{x}_n;\hat{\boldsymbol{x}}_{n|n-1},\boldsymbol{P}^{xx}_{n|n-1})\mathrm{d}\boldsymbol{x}_n\boldsymbol{H}^{\mathrm{T}}
\end{aligned} \tag{6-11}$$

可化简为

$$\boldsymbol{P}^{xz}_{n|n-1} = \boldsymbol{P}^{xx}_{n|n-1}\boldsymbol{H}^{\mathrm{T}} \tag{6-12}$$

6.3 线性卡尔曼滤波器

当动态过程和观测过程均为线性时,贝叶斯估计简化为 LKF。注意,预测方程可通过先验密度和后验密度的一般分布,由式(5-3)~式(5-7)直接得,因此卡尔曼滤波器适用于各种密度分布。只需要用到对动态方程和观测方程的线性度要求。但是,密度函数必须是已知的解析形式,这样才能采用某种方法计算

矩积分。在许多情况下，这已经被证明是一个必要条件。

表6－1给出了完整的LKF过程。

表6－1　线性卡尔曼滤波过程

步骤(1)	滤波器初始化	初始化和 $\hat{\boldsymbol{x}}_0$ 和 $\boldsymbol{P}_0^{xx}$
步骤(2)	状态矢量预测	$\hat{\boldsymbol{x}}_{n\mid n-1}=\boldsymbol{F}\hat{\boldsymbol{x}}_{n-1\mid n-1}+\boldsymbol{u}_n$， $\boldsymbol{P}_{n\mid n-1}^{xx}=\boldsymbol{F}\boldsymbol{P}_{n-1\mid n-1}^{xx}\boldsymbol{F}^{\mathrm{T}}+\boldsymbol{Q}$
步骤(3)	观测值预测	$\hat{z}_{n\mid n-1}=\boldsymbol{H}\hat{\boldsymbol{x}}_{n\mid n-1}$， $\boldsymbol{P}_{n\mid n-1}^{zz}=\boldsymbol{H}\boldsymbol{P}_{n\mid n-1}^{xx}\boldsymbol{H}^{\mathrm{T}}+\boldsymbol{R}$ $\boldsymbol{P}_{n\mid n-1}^{xz}=\boldsymbol{P}_{n\mid n-1}^{xx}\boldsymbol{H}^{\mathrm{T}}$
步骤(4)	卡尔曼滤波器更新	$\boldsymbol{K}_n\triangleq\boldsymbol{P}_{n\mid n-1}^{xz}(\boldsymbol{P}_{n\mid n-1}^{zz})^{-1}$ $\hat{\boldsymbol{x}}_{n\mid n}=\hat{\boldsymbol{x}}_{n\mid n-1}+\boldsymbol{K}_n(z_n^{\mathrm{o}}-\hat{z}_{n\mid n-1})$ $\boldsymbol{P}_{n\mid n}^{xx}=\boldsymbol{P}_{n\mid n-1}^{xx}-\boldsymbol{K}_n\boldsymbol{P}_{n\mid n-1}^{zz}\boldsymbol{K}_n^{\mathrm{T}}$
步骤(5)	存储结果	将 $\hat{\boldsymbol{x}}_{n\mid n}$ 和 $\boldsymbol{P}_{n\mid n}^{xx}$ 存储到跟踪文件中
步骤(6)	返回步骤(2)	

6.4　LKF在DIFAR浮标方位估计中的应用

LKF可以应用在单个DIFAR浮标中，利用包括方位角和方位变化率在内的方位观测值，来估计状态矢量。观测矢量定义如下

$$\boldsymbol{x}_n\triangleq[\theta_n,\dot{\theta}_n]^{\mathrm{T}}\tag{6－13}$$

式中：方位角 θ_n 为浮标原点至目标船舶的直线与 y 轴正向（指向正北方向）之间的夹角，规定 $-180°<\theta\leqslant180°$；$\dot{\theta}_n$ 为方位角变化率。对于简单问题，不需要控制变量 $\boldsymbol{u}_n$，动态转移方程由下式给出，即

$$\boldsymbol{x}_n=\boldsymbol{F}\boldsymbol{x}_{n-1}+\boldsymbol{v}_{n-1}\tag{6－14}$$

$$\boldsymbol{F}=\begin{bmatrix}1 & T\\0 & 1\end{bmatrix}\tag{6－15}$$

式中：$\boldsymbol{v}_{n-1}$ 为一个零均值高斯随机动态加速噪声过程，定义为 $\boldsymbol{v}_{n-1}\sim\mathcal{N}(\boldsymbol{0},\boldsymbol{Q})$；

$$\boldsymbol{Q}=q\begin{bmatrix}\dfrac{T^3}{3} & \dfrac{T^2}{2}\\[2mm]\dfrac{T^2}{2} & T\end{bmatrix}\tag{6－16}$$

$\boldsymbol{Q}$ 为连续噪声过程的动态噪声协方差；q 为方位角加速度噪声的方差，这里设为0.1。连续噪声模型下 $\boldsymbol{Q}$ 的完整推导超出了本书的范畴，感兴趣的读者可以在文

献[6]中找到。检验式(6-6)的状态协方差预测方程,得 q 的单位。注意下式,即

$$\begin{bmatrix} \sigma_\theta^2 & \sigma_{\theta\dot{\theta}}^2 \\ \sigma_{\theta\dot{\theta}}^2 & \sigma_{\dot{\theta}}^2 \end{bmatrix}_{n|n-1} = \boldsymbol{P}_{n|n-1}^{xx} = \boldsymbol{F}\boldsymbol{P}_{n-1|n-1}^{xx}\boldsymbol{F}^{\mathrm{T}} + \boldsymbol{Q}$$

$$= \boldsymbol{F}\begin{bmatrix} \sigma_\theta^2 & \sigma_{\theta\dot{\theta}}^2 \\ \sigma_{\theta\dot{\theta}}^2 & \sigma_{\dot{\theta}}^2 \end{bmatrix}_{n-1|n-1} \boldsymbol{F}^{\mathrm{T}}$$

$$+ q\begin{bmatrix} \frac{T^3}{3} & \frac{T^2}{2} \\ \frac{T^2}{2} & T \end{bmatrix} \tag{6-17}$$

q 的单位可推导出:由于 qT 的单位与 $\sigma_{\dot{\theta}}^2$ 或者(速度)^2 的单位相同,那么 q 的单位可表示为 v^2/t 或者 $a^2 \times t$。同时,q 是一个均值约为零的标准偏差,如果用动态噪声来代表恒定速度下的路径偏离,标准偏差需要考虑正、负偏差两种情况,因此是容许偏差的两倍。在这个实例中,允许的速度偏差为 $\sqrt{0.1}/2 = 0.16\text{rad/s}$。

对于单个 DIFAR 浮标,输出为 t_n 时刻的含噪方位观测值。线性观测方程可写为

$$z_n = \theta_n = \boldsymbol{H}\boldsymbol{x}_n + w_n$$

$$\omega_n \sim \mathcal{N}(0, \sigma_\theta^2) \tag{6-18}$$

$$\boldsymbol{H} = [1 \quad 0] \tag{6-19}$$

式中:σ_θ 是方位测量的标准偏差,为 DIFAR 浮标的恒定且时间无关的特征。

LKF 可以用初始化为

$$\hat{\boldsymbol{x}}_0 = \begin{bmatrix} \tilde{\theta}_0 \\ \dot{\theta}_0 \end{bmatrix}$$

式中:$\tilde{\theta}_0$ 为 DIFAR 的第一个方位观测值,可任意设置为 $\dot{\theta}_0 = -0.09\text{deg/s}$。令初始状态协方差矩阵为

$$\boldsymbol{P}_0^{xx} = \begin{bmatrix} \sigma_\theta^2 & 0 \\ 0 & \sigma_{\dot{\theta}}^2 \end{bmatrix} \tag{6-20}$$

式中,$\sigma_{\dot{\theta}} = -0.09\text{deg/s}$;$\sigma_\theta$ 为观测不确定性,单位为度。注意,在执行 LKF 过程中,方位角和方位变化率分别转换为弧度和弧度/秒,但线性实例中,这种转变并不必要。设 DIFAR 浮标的输入信号信噪比为 20dB,时宽带宽积为 90,这样可确保方位观测值为高斯分布。

图 6-1 给出了单 DIFAR 浮标方位跟踪 LKF 过程。注意在卡尔曼滤波器校正步骤,计算新息(有时也称作测量残差)的公式为

$$\eta_n = \tilde{\theta}_n - \hat{\theta}_{n|n-1} \tag{6-21}$$

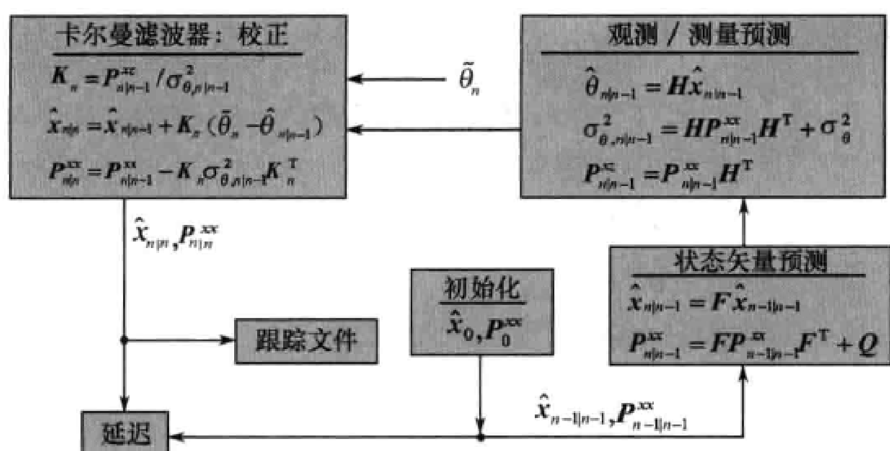


图 6 - 1　单 DIFAR 浮标 LKF 过程框图

有时，观测值和预测值正好落在 - 180/180 不连续点的两侧，这称为相位卷褶，必须特殊处理，否则跟踪滤波将失败。最简单的处理方式是计算 η_n，如果它大于 180°，则将 $\hat{\theta}_{n|n-1}$ 减去 360°，如果 η_n 小于 - 180°，$\hat{\theta}_{n|n-1}$ 加上 360°。如果上述条件均为假，不执行任何操作。

图 6 - 2 中可看到 DIFAR 浮标方位估计中的相位卷褶问题。如图所示，滤波器输出的估计方位用黑点表示，上面叠加的灰色线是从浮标到目标船舶真实方位。这种标绘方法称为方位时间记录（BTR）。使用 LKF 估计，选定滤波器更新率为 1s、方位加速度噪声为 0.1，可以得很好的方位估计。

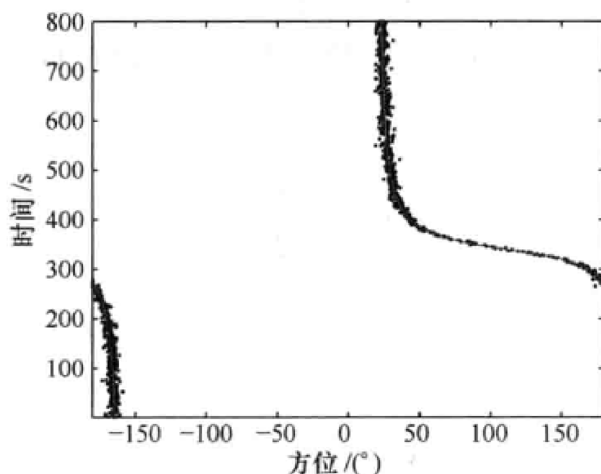


图 6 - 2　LKF 方位估计与单 DIFAR 浮标真实方位的比较

参考文献

[1] Kalman RE. A new approach to linear filtering and prediction problems. *ASME*, 1960,82(Series D):35 -40.

[2] Kalman RE, Bucy RS. New results in linear filtering and prediction theory. *ASME*, 1961, 83(Series D): 95 - 107.

[3] Papoulis A. Probability, Random Variables, and Stochastic Processes. McGRaw - Hill, 1965.

[4] Jazwinski AH. Stochastic Processes and Filtering Theory. Academic Press, 1970.

[5] Gelb A. *Applied Optimal Estimation.* The MIT Press, 1974.

[6] Bar Shalom Y, Li RX, Kirubarajan T. Estimation with Application to Tracking and Navigation: Theory, Algorithms and Software. Wiley, 2001.

第 7 章
线性化卡尔曼滤波器:扩展卡尔曼滤波器

对于适度非线性且平滑可微函数,可以采用解析方法将非线性方程线性化,这样线性卡尔曼滤波器就可以用来处理非线性动态方程和观测方程。为实现线性化,非线性函数 $\tilde{\boldsymbol{f}}(\boldsymbol{c})$ 可以展开为关于平均值的多维泰勒多项式 $\hat{\boldsymbol{c}}=\boldsymbol{0}$,得扩展卡尔曼滤波器[1,2]。为了更好地解释这种方法,并给后续章节做参考,首先推导标量扩展卡尔曼滤波器(EKF),然后将其推广至多维情况。

7.1 一维情况

对于一维情况,状态预测方程式(5-51)和式(5-52)变成

$$\hat{x}_{n|n-1}=\int\tilde{f}(c)\mathcal{N}(c;0,1)\mathrm{d}c+u_n \tag{7-1}$$

$$\sigma_{xx,n|n-1}^2=\int[\tilde{f}(c)+u_n-\hat{x}_{n|n-1}]^2\mathcal{N}(c;0,1)\mathrm{d}c+\sigma_{v,n}^2 \tag{7-2}$$

$$\tilde{f}(c)=f(\hat{x}_{n-1|n-1}+\sigma_{xx,n-1|n-1}c) \tag{7-3}$$

由式(5-55)至式(5-57)可知,对于一维情况,观测预测方程变成

$$\hat{z}_{n|n-1}=\int\tilde{h}(c)\mathcal{N}(c;0,1)\mathrm{d}c \tag{7-4}$$

$$\sigma_{zz,n|n-1}^2=\int\tilde{h}^2(c)\mathcal{N}(c;0,1)\mathrm{d}c-\hat{z}_{n|n-1}^2+\sigma_{w,n}^2 \tag{7-5}$$

$$\sigma_{xz,n|n-1}^2=\sigma_{xx,n|n-1}\int c\tilde{h}(c)\mathcal{N}(c;0,1)\mathrm{d}c \tag{7-6}$$

$$\tilde{h}(c)=h(\hat{x}_{n|n-1}+\sigma_{xx,n|n-1}c) \tag{7-7}$$

7.1.1 一维状态预测

式(2-38)为一般一维 M 次泰勒多项式。对于三阶一维泰勒多项式 $\tilde{f}(c)$,

关于零平均值 $\hat{c}$ 展开，则式(2－38) 变为

$$\tilde{f}(c) = \tilde{f}(0) + \left[\frac{\mathrm{d}}{\mathrm{d}c}\tilde{f}(c)\right]_{c=0} c + \frac{1}{2}\left[\frac{\mathrm{d}^2}{\mathrm{d}c^2}\tilde{f}(c)\right]_{c=0} c^2 + \frac{1}{6}\left[\frac{\mathrm{d}^3}{\mathrm{d}c^3}\tilde{f}(c)\right]_{c=0} c^3 \tag{7-8}$$

代入式(7－1)，得

$$\begin{aligned}\hat{x}_{n|n-1} = \tilde{f}(0)\int \mathcal{N}(c;\ 0,1)\mathrm{d}c + \\ \left[\frac{\mathrm{d}}{\mathrm{d}c}\tilde{f}(c)\right]_{c=0}\int c\mathcal{N}(c;\ 0,1)\mathrm{d}c + \\ \frac{1}{2}\left[\frac{\mathrm{d}^2}{\mathrm{d}c^2}\tilde{f}(c)\right]_{c=0}\int c^2\mathcal{N}(c;\ 0,1)\mathrm{d}c + \\ \frac{1}{6}\left[\frac{\mathrm{d}^3}{\mathrm{d}c^3}\tilde{f}(c)\right]_{c=0}\int c^3\mathcal{N}(c;\ 0,1)\mathrm{d}c + u_n\end{aligned} \tag{7-9}$$

使用一维矩方程(2－115) 计算积分，得状态矢量的一维预测估计为

$$\hat{x}_{n|n-1} = \tilde{f}(0) + \frac{1}{2}\left[\frac{\mathrm{d}^2}{\mathrm{d}c^2}\tilde{f}(c)\right]_{c=0} + u_n \tag{7-10}$$

注意，这个估计中 c 具有三阶精度，因为式(7－8) 中的三次项的积分为零。

7.1.2 一维状态估计误差方差预测

仅考虑式(7－2) 中的第一项

$$I(c) = \int[\tilde{f}(c) + u_n - \hat{x}_{n|n-1}]^2\mathcal{N}(c;\ 0,1)\mathrm{d}c \tag{7-11}$$

由式(7－1)，可进行如下推导，即

$$\begin{aligned}\hat{x}_{n|n-1} &= \int\tilde{f}(c)\mathcal{N}(c;\ 0,1)\mathrm{d}c + u_n \\ &= \int\tilde{f}(c)\mathcal{N}(c;\ 0,1)\mathrm{d}c + \tilde{f}(0) - \tilde{f}(0) + u_n \\ &= \tilde{f}(0) + \int[\tilde{f}(c) - \tilde{f}(0)]\mathcal{N}(c;\ 0,1)\mathrm{d}c + u_n\end{aligned} \tag{7-12}$$

因此式(7－11) 变成

$$\begin{aligned}I(c) &= \int\left\{[\tilde{f}(c) - \tilde{f}(0)] - \int[\tilde{f}(\tau) - \tilde{f}(0)]\mathcal{N}(\tau;\ 0,1)\mathrm{d}\tau\right\}^2\mathcal{N}(c;\ 0,1)\mathrm{d}c \\ &= \int[\tilde{f}(c) - \tilde{f}(0)]^2\mathcal{N}(c;\ 0,1)\mathrm{d}c - \\ &\quad \left\{\int[\tilde{f}(c) - \tilde{f}(0)]\mathcal{N}(c;\ 0,1)\mathrm{d}c\right\}^2\end{aligned} \tag{7-13}$$

将泰勒多项式(7－8) 保留至二次项，得

$$[\tilde{f}(c) - \tilde{f}(0)] = \left[\frac{\mathrm{d}}{\mathrm{d}c}\tilde{f}(c)\right]_{c=0} c + \frac{1}{2}\left[\frac{\mathrm{d}^2}{\mathrm{d}c^2}\tilde{f}(c)\right]_{c=0} c^2 \tag{7-14}$$

$$[\tilde{f}(c)-\tilde{f}(0)]^2=\left[\frac{d}{dc}\tilde{f}(c)\right]_{c=0}^2 c^2+\left[\frac{d}{dc}\tilde{f}(c)\right]_{c=0}\left[\frac{d^2}{dc^2}\tilde{f}(c)\right]_{c=0}c^3$$
$$+\frac{1}{4}\left[\frac{d^2}{dc^2}\tilde{f}(c)\right]_{c=0}^2 c^4 \qquad (7-15)$$

将式(7－14)、式(7－15)代入式(7－13)，利用一维矩方程式(2－115)计算合成积分，得

$$I(c)=\left[\frac{d}{dc}\tilde{f}(c)\right]_{c=0}^2+\frac{1}{2}\left[\frac{d^2}{dc^2}\tilde{f}(c)\right]_{c=0}^2 \qquad (7-16)$$

因此，状态误差方差可写为

$$\sigma_{xx,n|n-1}^2=\left[\frac{d}{dc}\tilde{f}(c)\right]_{c=0}^2+\frac{1}{2}\left[\frac{d^2}{dc^2}\tilde{f}(c)\right]_{c=0}^2+\sigma_{v,n}^2 \qquad (7-17)$$

由式(7－15)可以看出，估计精度可达到 c 的5次幂。

7.1.3 一维观测预测方程

按照和上面两节相同的步骤，对 $\hat{z}_{n|n-1}$ 和 $\sigma_{zz,n|n-1}^2$ 进行推导，得

$$\hat{z}_{n|n-1}=\tilde{h}(0)+\frac{1}{2}\left[\frac{d^2}{dc^2}\tilde{h}(c)\right]_{c=0} \qquad (7-18)$$

$$\sigma_{zz,n|n-1}^2=\left[\frac{d}{dc}\tilde{h}(c)\right]_{c=0}^2+\frac{1}{2}\left[\frac{d^2}{dc^2}\tilde{h}(c)\right]_{c=0}^2+\sigma_{w,n}^2 \qquad (7-19)$$

互相关系数为

$$\begin{aligned}\sigma_{xz,n|n-1}^2&=\sigma_{xx,n|n-1}\int c\tilde{h}(c)\mathcal{N}(c;0,1)dc\\&=\sigma_{xx,n|n-1}\int c\left\{\tilde{h}(0)+\left[\frac{d}{dc}\tilde{h}(c)\right]_{c=0}c+\right.\\&\qquad\left.\frac{1}{2}\left[\frac{d^2}{dc^2}\tilde{h}(c)\right]_{c=0}c^2\right\}\mathcal{N}(c;0,1)dc\\&=\sigma_{xx,n|n-1}\left[\frac{d}{dc}\tilde{h}(c)\right]_{c=0}\end{aligned} \qquad (7-20)$$

对互相关系数的估计可达到 c 的3次幂。

7.1.4 一维预测方程的变换

根据仿射变换的定义(5－46)，利用微分链式法则，有

$$\frac{d}{dc}=\frac{dx_{n-1}}{dc}\frac{d}{dx_{n-1}}=\sigma_{xx,n-1|n-1}\frac{d}{dx_{n-1}} \qquad (7-21)$$

$$\frac{d^2}{dc^2}=\left(\frac{dx_{n-1}}{dc}\right)^2\frac{d^2}{dx_{n-1}^2}$$

$$= \sigma_{xx,n-1|n-1}^{2} \frac{\mathrm{d}^2}{\mathrm{d}x_{n-1}^2} \tag{7-22}$$

由式(7 - 3),有 $\tilde{f}(0) = f(\hat{x}_{n-1|n-1})$;由式(5 - 46) 有 $c = 0 \rightarrow x_{n-1} = \hat{x}_{n-1|n-1}$。这样式(7 - 10) 变成

$$\hat{x}_{n|n-1} = f(\hat{x}_{n-1|n-1}) + \frac{1}{2}\sigma_{xx,n-1|n-1}^{2}\left[\frac{\mathrm{d}^2}{\mathrm{d}x_{n-1}^2}f(x_{n-1})\right]_{x_{n-1}=\hat{x}_{n-1|n-1}} + u_n \tag{7-23}$$

用同样方法,将式(7 - 17) 变成

$$\sigma_{xx,n|n-1}^{2} = \sigma_{xx,n-1|n-1}^{2}\left[\frac{\mathrm{d}}{\mathrm{d}x_{n-1}}f(x_{n-1})\right]_{x_{n-1}=\hat{x}_{n-1|n-1}}^{2} + \frac{1}{2}\sigma_{xx,n-1|n-1}^{4}\left[\frac{\mathrm{d}^2}{\mathrm{d}x_{n-1}^2}f(x_{n-1})\right]_{x_{n-1}=-\hat{x}_{n-1|n-1}}^{2} + \sigma_{v,n}^{2} \tag{7-24}$$

对于观测方程,利用式(5 - 46) 仿射变换,而不是式(7 - 21)、式(7 - 22),链式法则变成

$$\frac{\mathrm{d}}{\mathrm{d}c} = \frac{\mathrm{d}x_n}{\mathrm{d}c}\frac{\mathrm{d}}{\mathrm{d}x_n} = \sigma_{xx,n|n-1}\frac{\mathrm{d}}{\mathrm{d}x_n} \tag{7-25}$$

$$\frac{\mathrm{d}^2}{\mathrm{d}c^2} = \sigma_{xx,n|n-1}^{2}\frac{\mathrm{d}^2}{\mathrm{d}x_n^2} \tag{7-26}$$

这样,式(7 - 18) ~ 式(7 - 20) 变换为

$$\hat{z}_{n|n-1} = h(\hat{x}_{n|n-1}) + \frac{1}{2}\sigma_{xx,n|n-1}^{2}\left[\frac{\mathrm{d}^2}{\mathrm{d}x_n^2}h(x_{n-1})\right]_{x_n=\hat{x}_{n|n-1}} \tag{7-27}$$

$$\sigma_{zz,n|n-1}^{2} = \sigma_{xx,n|n-1}^{2}\left[\frac{\mathrm{d}}{\mathrm{d}x_n}h(x_n)\right]_{x_n=\hat{x}_{n|n-1}}^{2} + \frac{1}{2}\sigma_{xx,n|n-1}^{4}\left[\frac{\mathrm{d}^2}{\mathrm{d}x_n^2}f(x_n)\right]_{x_n=\hat{x}_{n|n-1}}^{2} + \sigma_{w,n}^{2} \tag{7-28}$$

$$\sigma_{xz,n|n-1}^{2} = \sigma_{xx,n|n-1}^{2}\left[\frac{\mathrm{d}}{\mathrm{d}x_n}h(x_n)\right]_{x_n=\hat{x}_{n|n-1}} \tag{7-29}$$

7.1.5　一维线性化 EKF 过程

上面提出的 EKF 中,包含比线性次数高的项。标准线性 EKF 可以通过截断,即只保留线性项来得。这样得的 EKF 和第 6 章中给出的线性卡尔曼滤波器具有相同形式。一维线性化 EKF 过程如表 7 - 1 所列。注意,对于这种一维情况,有 $\rho_{xz,n|n-1} = \rho_{zx,n|n=1}$。在更新步骤中,$z_n^0$ 为 t_n 时刻的观测值。

表 7-1　一维扩展卡尔曼滤波器过程

步骤(1)	滤波器初始化	初始化 $\hat{x}_0$ 和 $\boldsymbol{\sigma}_{xx,0}^2$
步骤(2)	状态矢量预测	$\hat{x}_{n\mid n-1}=f(\hat{x}_{n-1\mid n-1})+u_n$ $\sigma_{xx,n\mid n-1}^2=\sigma_{xx,n-1\mid n-1}^2\times\left[\frac{\mathrm{d}}{\mathrm{d}x_{n-1}}f(x_{n-1})\right]_{x_{n-1}=\hat{x}_{n-1\mid n-1}}^2+\sigma_{v,n}^2$
步骤(3)	观测值预测	$\hat{z}_{n\mid n-1}=h(\hat{x}_{n\mid n-1})$ $\sigma_{zz,n\mid n-1}^2=\sigma_{xx,n\mid n-1}^2\left[\frac{\mathrm{d}}{\mathrm{d}x_n}h(x_n)\right]_{x_n=\hat{x}_{n\mid n-1}}^2+\sigma_{w,n}^2$ $\sigma_{xz,n\mid n-1}^2=\sigma_{xx,n\mid n-1}^2\left[\frac{\mathrm{d}}{\mathrm{d}x_n}h(x_n)\right]_{x_n=\hat{x}_{n\mid n-1}}$
步骤(4)	卡尔曼滤波器更新	$K_n=\sigma_{xz,n\mid n-1}^2/\sigma_{zz,n\mid n-1}^2$ $\hat{x}_{n\mid n}=\hat{x}_{n\mid n-1}+K_n(z_n^o-\hat{z}_{n\mid n-1})$ $\sigma_{xx,n\mid n}^2=\sigma_{xx,n\mid n-1}^2-K_n^2\sigma_{zz,n\mid n-1}^2$
步骤(5)	存储结果	将 $\hat{\boldsymbol{x}}_{n\mid n}$ 和 $\boldsymbol{\sigma}_{xx,n\mid n}^2$ 存储到航迹文件中
步骤(6)	返回步骤(2)	

7.2　多维情况

对于多维参数的多维函数，$\tilde{f}(\boldsymbol{c})$ 的状态预测积分由式(5-51)～式(5-57)给出。由式(6-3)、式(7-23)很容易看出，在线性卡尔曼滤波器和一维EKF中，确定性控制方程 $\boldsymbol{u}_n$ 对于状态预测值是加性的，在其他滤波方程中不是这样的。所以，后面将忽略 $\boldsymbol{u}_n$，如果实例研究需要，再把它加回去。

式(2-63)给出一般多维泰勒多项式，关于 $\boldsymbol{c}$ 的零均值，进行 $\tilde{\boldsymbol{f}}(\boldsymbol{c})$ 的二次泰勒多项式展开，将式(2-69)代入式(2-63)，得

$$\begin{aligned}\tilde{\boldsymbol{f}}(\boldsymbol{c}) &= \sum_{i=0}^{2}\frac{1}{i!}D_c^i\tilde{\boldsymbol{f}}(\boldsymbol{c})\\ &= \tilde{\boldsymbol{f}}(\boldsymbol{0})+\sum_{i=1}^{n_x}c_i\left[\frac{\partial}{\partial c_i}\tilde{\boldsymbol{f}}(\boldsymbol{c})\right]_{\boldsymbol{c}=\boldsymbol{0}}+\\ &\quad\frac{1}{2}\sum_{\substack{i_1=0\\ i_1+\cdots+i_{n_x}=2}}^{2}\cdots\sum_{i_{n_x}=0}^{2}\binom{2}{i_1\cdots i_{n_x}}c_1^{i_1}\cdots c_{n_x}^{i_{n_x}}\cdot\\ &\quad\left[\frac{\partial^2}{\partial c_1^{i_1}\cdots\partial c_{n_x}^{i_{n_x}}}\tilde{\boldsymbol{f}}(\boldsymbol{c})\right]_{\boldsymbol{c}=\boldsymbol{0}}\end{aligned}\tag{7-30}$$

因此,式(7 - 30)可以重写为

$$\tilde{\boldsymbol{f}}(\boldsymbol{c}) = \tilde{\boldsymbol{f}}(\boldsymbol{0}) + \left[\sum_{i=1}^{n_x} \frac{\partial}{\partial c_i}\tilde{\boldsymbol{f}}(\boldsymbol{c})\right]_{c=0} c_i + \frac{1}{2}\left[\sum_{i=1}^{n_x} \frac{\partial^2}{\partial c_i^2}\tilde{\boldsymbol{f}}(\boldsymbol{c})\right]_{c=0} c_i^2 + \left[\sum_{i=1}^{n_x}\sum_{\substack{j=1\\ j\neq i}}^{n_x} \frac{\partial^2}{\partial c_i \partial c_j}\tilde{\boldsymbol{f}}(\boldsymbol{c})\right]_{c=0} c_i c_j \tag{7-31}$$

7.2.1 状态预测方程

由式(5 - 51),利用泰勒多项式(7 - 31),可以将预测一阶矩表示为

$$\hat{\boldsymbol{x}}_{n|n-1} = \tilde{\boldsymbol{f}}(\boldsymbol{0})\int \mathcal{N}(\boldsymbol{c};\boldsymbol{0},\boldsymbol{I})\,\mathrm{d}\boldsymbol{c} + \left[\sum_{i=1}^{n_x} \frac{\partial}{\partial c_i}\tilde{\boldsymbol{f}}(\boldsymbol{c})\right]_{c=0}\int c_i \mathcal{N}(\boldsymbol{c};\boldsymbol{0},\boldsymbol{I})\,\mathrm{d}\boldsymbol{c} + \frac{1}{2}\left[\sum_{i=1}^{n_x} \frac{\partial}{\partial c_i^2}\tilde{\boldsymbol{f}}(\boldsymbol{c})\right]_{c=0}\int c_i^2\, \mathcal{N}(\boldsymbol{c};\boldsymbol{0},\boldsymbol{I})\,\mathrm{d}\boldsymbol{c} + \left[\sum_{i=1}^{n_x}\sum_{\substack{j=1\\ j\neq i}}^{n_x} \frac{\partial^2}{\partial c_i \partial c_j}\tilde{\boldsymbol{f}}(\boldsymbol{c})\right]_{c=0}\int c_i c_j \mathcal{N}(\boldsymbol{c};\boldsymbol{0},\boldsymbol{I})\,\mathrm{d}\boldsymbol{c} \tag{7-32}$$

利用矩方程(2 - 116) ~ 式(2 - 121),式(7 - 32)可简化为

$$\hat{\boldsymbol{x}}_{n|n-1} = \tilde{\boldsymbol{f}}(\boldsymbol{0}) + \frac{1}{2}\left[\sum_{i=1}^{n_x} \frac{\partial^2}{\partial c_i^2}\tilde{\boldsymbol{f}}(\boldsymbol{c})\right]_{c=0} \tag{7-33}$$

如果式(7 - 31)中包含三次项,在式(7 - 32)中它们的和将为零,因此式(7 - 33)对于三次是有效的。

7.2.2 状态协方差预测方程

由式(5 - 52),得协方差预测方程为

$$\boldsymbol{P}_{n|n-1}^{xx} = \int [\tilde{\boldsymbol{f}}(\boldsymbol{c}) - \hat{\boldsymbol{x}}_{n|n-1}][\tilde{\boldsymbol{f}}(\boldsymbol{c}) - \hat{\boldsymbol{x}}_{n|n-1}]^{\mathrm{T}} \mathcal{N}(\boldsymbol{c};\boldsymbol{0},\boldsymbol{I})\,\mathrm{d}\boldsymbol{c} + \boldsymbol{Q} \tag{7-34}$$

注意,通过加减 $\tilde{\boldsymbol{f}}(\boldsymbol{0})$,式(5 - 51)可重写为

$$\hat{\boldsymbol{x}}_{n|n-1} = \tilde{\boldsymbol{f}}(\boldsymbol{0}) + \int [\tilde{\boldsymbol{f}}(\boldsymbol{c}) - \tilde{\boldsymbol{f}}(\boldsymbol{0})] \mathcal{N}(\boldsymbol{c};\boldsymbol{0},\boldsymbol{I})\,\mathrm{d}\boldsymbol{c} \tag{7-35}$$

用式(7 - 35)替换 $\hat{\boldsymbol{x}}_{n|n-1}$,式(7 - 34)变为

$$\boldsymbol{P}_{n|n-1}^{xx} = \int [\tilde{\boldsymbol{f}}(\boldsymbol{c}) - \tilde{\boldsymbol{f}}(\boldsymbol{0})][\tilde{\boldsymbol{f}}(\boldsymbol{c}) - \tilde{\boldsymbol{f}}(\boldsymbol{0})]^{\mathrm{T}} \mathcal{N}(\boldsymbol{c};\boldsymbol{0},\boldsymbol{I})\,\mathrm{d}\boldsymbol{c} - \left\{\int [\tilde{\boldsymbol{f}}(\boldsymbol{c}) - \tilde{\boldsymbol{f}}(\boldsymbol{0})] \mathcal{N}(\boldsymbol{c};\boldsymbol{0},\boldsymbol{I})\,\mathrm{d}\boldsymbol{c}\right\} \left\{\int [\tilde{\boldsymbol{f}}(\boldsymbol{c}) - \tilde{\boldsymbol{f}}(\boldsymbol{0})]^{\mathrm{T}} \mathcal{N}(\boldsymbol{c};\boldsymbol{0},\boldsymbol{I})\,\mathrm{d}\boldsymbol{c}\right\} + \boldsymbol{Q} \tag{7-36}$$

重写式(7 - 31)的二次泰勒多项式,得

$$\tilde{\boldsymbol{f}}(\boldsymbol{c})-\tilde{\boldsymbol{f}}(\boldsymbol{0})=\left[\sum_{i=1}^{n_x}\frac{\partial}{\partial c_i}\tilde{\boldsymbol{f}}(\boldsymbol{c})\right]_{\boldsymbol{c}=\boldsymbol{0}}c_i+\frac{1}{2}\left[\sum_{i=1}^{n_x}\frac{\partial^2}{\partial c_i^2}\tilde{\boldsymbol{f}}(\boldsymbol{c})\right]_{\boldsymbol{c}=\boldsymbol{0}}c_i^2+$$

$$\left[\sum_{i=1}^{n_x}\sum_{\substack{j=1\\j\neq i}}^{n_x}\frac{\partial^2}{\partial c_i\partial c_j}\tilde{\boldsymbol{f}}(\boldsymbol{c})\right]_{\boldsymbol{c}=\boldsymbol{0}}c_ic_j \tag{7-37}$$

利用矩方程式(2 - 116) ~ 式(2 - 121),则式(7 - 36) 中第二个积分项变成

$$\int[\tilde{\boldsymbol{f}}(\boldsymbol{c})-\tilde{\boldsymbol{f}}(\boldsymbol{0})]\mathcal{N}(\boldsymbol{c};\boldsymbol{0},\boldsymbol{I})\mathrm{d}\boldsymbol{c}=\frac{1}{2}\left[\sum_{i=1}^{n_x}\frac{\partial^2}{\partial c_i^2}\tilde{\boldsymbol{f}}(\boldsymbol{c})\right]_{\boldsymbol{c}=\boldsymbol{0}} \tag{7-38}$$

同样,由二阶泰勒多项式(7 - 31),可以有

$$\begin{aligned}&[\tilde{\boldsymbol{f}}(\boldsymbol{c})-\tilde{\boldsymbol{f}}(\boldsymbol{0})][\tilde{\boldsymbol{f}}(\boldsymbol{c})-\tilde{\boldsymbol{f}}(\boldsymbol{0})]^{\mathrm{T}}\\
&=\left\{\left[\sum_{i=1}^{n_x}\frac{\partial}{\partial c_i}\tilde{\boldsymbol{f}}(\boldsymbol{c})\right]_{\boldsymbol{c}=\boldsymbol{0}}c_i+\frac{1}{2}\left[\sum_{i=1}^{n_x}\frac{\partial^2}{\partial c_i^2}\tilde{\boldsymbol{f}}(\boldsymbol{c})\right]_{\boldsymbol{c}=\boldsymbol{0}}c_i^2+\right.\\
&\left.\left[\sum_{i=1}^{n_x}\sum_{\substack{j=1\\j\neq i}}^{n_x}\frac{\partial^2}{\partial c_i\partial c_j}\tilde{\boldsymbol{f}}(\boldsymbol{c})\right]_{\boldsymbol{c}=\boldsymbol{0}}c_ic_j\right\}\cdot\\
&\left\{\left[\sum_{i=1}^{n_x}\frac{\partial}{\partial c_i}\tilde{\boldsymbol{f}}^{\mathrm{T}}(\boldsymbol{c})\right]_{\boldsymbol{c}=\boldsymbol{0}}c_i+\frac{1}{2}\left[\sum_{i=1}^{n_x}\frac{\partial^2}{\partial c_i^2}\tilde{\boldsymbol{f}}^{\mathrm{T}}(\boldsymbol{c})\right]_{\boldsymbol{c}=\boldsymbol{0}}c_i^2+\right.\\
&\left.\left[\sum_{i=1}^{n_x}\sum_{\substack{j=1\\j\neq i}}^{n_x}\frac{\partial^2}{\partial c_i\partial c_j}\tilde{\boldsymbol{f}}^{\mathrm{T}}(\boldsymbol{c})\right]_{\boldsymbol{c}=\boldsymbol{0}}c_ic_j\right\}\end{aligned} \tag{7-39}$$

逐项进行外积运算,将结果代入式(7 - 36) 的第一个积分项,使用矩方程式(2 - 116) ~ 式(2 - 121),得

$$\begin{aligned}&\int[\tilde{\boldsymbol{f}}(\boldsymbol{c})-\tilde{\boldsymbol{f}}(\boldsymbol{0})][\tilde{\boldsymbol{f}}(\boldsymbol{c})-\tilde{\boldsymbol{f}}(\boldsymbol{0})]^{\mathrm{T}}\mathcal{N}(\boldsymbol{c};\boldsymbol{0},\boldsymbol{I})\mathrm{d}\boldsymbol{c}\\
&=\left[\sum_{i=1}^{n_x}\frac{\partial}{\partial c_i}\tilde{\boldsymbol{f}}(\boldsymbol{c})\right]_{\boldsymbol{c}=\boldsymbol{0}}\left[\sum_{i=1}^{n_x}\frac{\partial}{\partial c_i}\tilde{\boldsymbol{f}}^{\mathrm{T}}(\boldsymbol{c})\right]_{\boldsymbol{c}=\boldsymbol{0}}+\\
&\frac{3}{4}\left[\sum_{i=1}^{n_c}\frac{\partial^2}{\partial c_i^2}\tilde{\boldsymbol{f}}(\boldsymbol{c})\right]_{\boldsymbol{c}=\boldsymbol{0}}\left[\sum_{i=1}^{n_c}\frac{\partial^2}{\partial c_i^2}\tilde{\boldsymbol{f}}^{\mathrm{T}}(\boldsymbol{c})\right]_{\boldsymbol{c}=\boldsymbol{0}}+\\
&\left[\sum_{i=1}^{n_x}\sum_{\substack{j=1\\j\neq i}}^{n_x}\frac{\partial^2}{\partial c_i\partial c_j}\tilde{\boldsymbol{f}}(\boldsymbol{c})\right]_{\boldsymbol{c}=\boldsymbol{0}}\left[\sum_{i=1}^{n_x}\sum_{\substack{j=1\\j\neq i}}^{n_x}\frac{\partial^2}{\partial c_i\partial c_j}\tilde{\boldsymbol{f}}^{\mathrm{T}}(\boldsymbol{c})\right]_{\boldsymbol{c}=\boldsymbol{0}}\end{aligned} \tag{7-40}$$

式(7 - 36) 则变为

$$\begin{aligned}\boldsymbol{P}_{n|n-1}^{xx}=&\left[\sum_{i=1}^{n_x}\frac{\partial}{\partial c_i}\tilde{\boldsymbol{f}}(\boldsymbol{c})\right]_{\boldsymbol{c}=\boldsymbol{0}}\left[\sum_{i=1}^{n_x}\frac{\partial}{\partial c_i}\tilde{\boldsymbol{f}}^{\mathrm{T}}(\boldsymbol{c})\right]_{\boldsymbol{c}=\boldsymbol{0}}+\\
&\frac{1}{2}\left[\sum_{i=1}^{n_x}\frac{\partial^2}{\partial c_i^2}\tilde{\boldsymbol{f}}(\boldsymbol{c})\right]_{\boldsymbol{c}=\boldsymbol{0}}\left[\sum_{i=1}^{n_x}\frac{\partial^2}{\partial c_i^2}\tilde{\boldsymbol{f}}^{\mathrm{T}}(\boldsymbol{c})\right]_{\boldsymbol{c}=\boldsymbol{0}}+\end{aligned}$$

$$\left[\sum_{i=1}^{n_x}\sum_{\substack{j=1\\j\neq i}}^{n_x}\frac{\partial^2}{\partial c_i\partial c_j}\tilde{\boldsymbol{f}}(\boldsymbol{c})\right]_{\boldsymbol{c}=\boldsymbol{0}}\left[\sum_{i=1}^{n_x}\sum_{\substack{j=1\\j\neq i}}^{n_x}\frac{\partial^2}{\partial c_i\partial c_j}\tilde{\boldsymbol{f}}^{\mathrm{T}}(\boldsymbol{c})\right]_{\boldsymbol{c}=\boldsymbol{0}}+\boldsymbol{Q} \tag{7-41}$$

如果保留所有项,状态误差协方差矩阵一直到五阶均有效,因为第五阶项的积分为零。

7.2.3 观测预测方程

用同样的方式,预测估计 $\hat{z}_{n|n-1}$ 和 $\boldsymbol{P}^{zz}_{n|n-1}$ 可表示为

$$\hat{z}_{n|n-1}=\tilde{\boldsymbol{h}}(\boldsymbol{0})+\frac{1}{2}\left[\sum_{i=1}^{n_x}\frac{\partial^2}{\partial c_i^2}\tilde{\boldsymbol{h}}(\boldsymbol{c})\right]_{\boldsymbol{c}=\boldsymbol{0}} \tag{7-42}$$

$$\begin{aligned}\boldsymbol{P}^{zz}_{n|n-1}=&\left[\sum_{i=1}^{n_x}\frac{\partial}{\partial c_i}\tilde{\boldsymbol{h}}(\boldsymbol{c})\right]_{\boldsymbol{c}=\boldsymbol{0}}\left[\sum_{i=1}^{n_x}\frac{\partial}{\partial c_i}\tilde{\boldsymbol{h}}^{\mathrm{T}}(\boldsymbol{c})\right]_{\boldsymbol{c}=\boldsymbol{0}}+\\&\frac{1}{2}\left[\sum_{i=1}^{n_x}\frac{\partial^2}{\partial c_i^2}\tilde{\boldsymbol{h}}(\boldsymbol{c})\right]_{\boldsymbol{c}=\boldsymbol{0}}\left[\sum_{i=1}^{n_x}\frac{\partial^2}{\partial c_i^2}\tilde{\boldsymbol{h}}^{\mathrm{T}}(\boldsymbol{c})\right]_{\boldsymbol{c}=\boldsymbol{0}}+\\&\left[\sum_{i=1}^{n_x}\sum_{\substack{j=1\\j\neq i}}^{n_x}\frac{\partial^2}{\partial c_i\partial c_j}\tilde{\boldsymbol{h}}(\boldsymbol{c})\right]_{\boldsymbol{c}=\boldsymbol{0}}\left[\sum_{i=1}^{n_x}\sum_{\substack{j=1\\j\neq i}}^{n_x}\frac{\partial^2}{\partial c_i\partial c_j}\tilde{\boldsymbol{h}}^{\mathrm{T}}(\boldsymbol{c})\right]_{\boldsymbol{c}=\boldsymbol{0}}+\boldsymbol{R}\end{aligned} \tag{7-43}$$

最后,$\boldsymbol{P}^{xz}_{n|n-1}$ 由式(5 - 57) 给出,可写为

$$\boldsymbol{P}^{xz}_{n|n-1}=\boldsymbol{D}_{n|n-1}\int\boldsymbol{c}\,\tilde{\boldsymbol{h}}_n^{\mathrm{T}}(\boldsymbol{c})\mathcal{N}(\boldsymbol{c};\boldsymbol{0},\boldsymbol{I})\,\mathrm{d}\boldsymbol{c} \tag{7-44}$$

利用式(7 - 31) 将 $\boldsymbol{h}_n(\boldsymbol{c})$ 展开为泰勒多项式,则式(7 - 44) 可表示为

$$\begin{aligned}\boldsymbol{P}^{xz}_{n|n-1}=&\boldsymbol{D}_{n|n-1}\tilde{\boldsymbol{h}}^{\mathrm{T}}(\boldsymbol{0})\int\boldsymbol{c}\mathcal{N}(\boldsymbol{c};\boldsymbol{0},\boldsymbol{I})\,\mathrm{d}\boldsymbol{c}+\\&\boldsymbol{D}_{n|n-1}\left[\sum_{i=1}^{n_x}\frac{\partial}{\partial c_i}\tilde{\boldsymbol{h}}^{\mathrm{T}}(\boldsymbol{c})\right]_{\boldsymbol{c}=\boldsymbol{0}}\int\boldsymbol{c}c_i\mathcal{N}(\boldsymbol{c};\boldsymbol{0},\boldsymbol{I})\,\mathrm{d}\boldsymbol{c}+\\&\boldsymbol{D}_{n|n-1}\left[\sum_{i=1}^{n_x}\sum_{\substack{j=1\\j\neq i}}^{n_x}\frac{\partial}{\partial c_i}\frac{\partial}{\partial c_j}\tilde{\boldsymbol{h}}^{\mathrm{T}}(\boldsymbol{c})\right]_{\boldsymbol{c}=\boldsymbol{0}}\cdot\\&\int\boldsymbol{c}c_ic_j\mathcal{N}(\boldsymbol{c};\boldsymbol{0},\boldsymbol{I})\,\mathrm{d}\boldsymbol{c}\end{aligned} \tag{7-45}$$

计算积分结果

$$\boldsymbol{P}^{xz}_{n|n-1}=\boldsymbol{D}_{n|n-1}\left[\sum_{i=1}^{n_x}\frac{\partial}{\partial c_i}\tilde{\boldsymbol{h}}_i(\boldsymbol{c})\right]_{\boldsymbol{c}=\boldsymbol{0}} \tag{7-46}$$

注意:结果具有三阶精度,因为 $\boldsymbol{c}$ 的三次项为零。

7.2.4 多维预测方程的变换

由式(2 - 5) 矩阵的迹的定义,有

$$\sum_{i=1}^{n_x} \frac{\partial^2}{\partial c_i^2} \triangleq \mathrm{tr}\{ \boldsymbol{\nabla}_c \boldsymbol{\nabla}_c^{\mathrm{T}} \} \tag{7-47}$$

利用矢量链式法则

$$\boldsymbol{\nabla}_c = [\boldsymbol{\nabla}_c \boldsymbol{x}^{\mathrm{T}}] \boldsymbol{\nabla}_x \tag{7-48}$$

利用式(2 - 107),得

$$\boldsymbol{\nabla}_c \boldsymbol{x}^{\mathrm{T}} = \boldsymbol{\nabla}_c [\hat{x} + \boldsymbol{D}\boldsymbol{c}]^{\mathrm{T}} = \boldsymbol{D}^{\mathrm{T}} \tag{7-49}$$

因此,式(7 - 47) 变成

$$\sum_{i=1}^{n_x} \frac{\partial^2}{\partial c_i^2} = \mathrm{tr}\{ \boldsymbol{D}^{\mathrm{T}} \boldsymbol{\nabla}_x \boldsymbol{\nabla}_x^{\mathrm{T}} \boldsymbol{D} \} \tag{7-50}$$

由于迹具有循环置换不变性,式(7 - 50) 可重写为

$$\sum_{i=1}^{n_x} \frac{\partial^2}{\partial c_i^2} = \mathrm{tr}\{ \boldsymbol{D}\boldsymbol{D}^{\mathrm{T}} \boldsymbol{\nabla}_x \boldsymbol{\nabla}_x^{\mathrm{T}} \} = \mathrm{tr}\{ \boldsymbol{P}^{xx} \hat{\boldsymbol{G}}_{n-1|n-1} \} \tag{7-51}$$

这里 $\hat{\boldsymbol{G}}_{n-1|n-1}$ 定义为 Hessian 矩阵,则有

$$\hat{\boldsymbol{G}}_{n-1|n-1} \triangleq \boldsymbol{\nabla}_{x_{n-1}} \boldsymbol{\nabla}_{x_{n-1}}^{\mathrm{T}} = \begin{bmatrix} \frac{\partial^2}{\partial x_{n-1,1}^2} & \cdots & \frac{\partial}{\partial x_{n-1,1}} \frac{\partial}{\partial x_{n-1,n_x}} \\ \vdots & \ddots & \vdots \\ \frac{\partial}{\partial x_{n-1,n_x}} & \frac{\partial}{\partial x_{n-1,1}} & \cdots \frac{\partial^2}{\partial x_{n-1,n_x}^2} \end{bmatrix}_{x_{n-1}=\hat{x}_{n-1|n-1}} \tag{7-52}$$

式中:$\partial/\partial x_{n-1,i}$ 表示对于矢量 $\boldsymbol{x}_{n-1}$ 的第 i 个分量求偏导。

对于状态矢量预测,由式(5 - 46) 得,当 $\boldsymbol{c} = \boldsymbol{0}$ 时 $\boldsymbol{x} = \hat{\boldsymbol{x}}_{n-1|n-1}$;由式(5 - 60) 得 $\tilde{\boldsymbol{f}}(\boldsymbol{0}) = \boldsymbol{f}(\tilde{\boldsymbol{x}}_{n-1|n-1})$,这样式(7 - 33) 可写为

$$\hat{\boldsymbol{x}}_{n|n-1} = \boldsymbol{f}(\hat{\boldsymbol{x}}_{n-1|n-1}) + \frac{1}{2}\mathrm{tr}\{ \boldsymbol{P}_{n-1|n-1}^{xx} \hat{\boldsymbol{G}}_{n-1|n-1} \} \boldsymbol{f}(\boldsymbol{x}_{n-1}) \Big|_{x_{n-1}=\hat{x}_{n-1|n-1}} \tag{7-53}$$

对 $\boldsymbol{f}(\boldsymbol{x}_{n-1})$ 中的每一个元素进行迹 $\{ \boldsymbol{P}_{n-1|n-1}^{xx} \hat{\boldsymbol{G}}_{n-1|n-1} \}$ 运算。

遵循同样的逻辑,式(7 - 41) 可变换为

$$\boldsymbol{P}_{n|n-1}^{xx} = \boldsymbol{Q} + \hat{\boldsymbol{F}}_{n-1|n-1} \boldsymbol{P}_{n-1|n-1}^{xx} \hat{\boldsymbol{F}}_{n-1|n-1}^{\mathrm{T}} + \frac{1}{2}[\mathrm{tr}\{ \boldsymbol{P}_{n-1|n-1}^{xx} \hat{\boldsymbol{G}}_{n-1|n-1} \} \boldsymbol{f}(\boldsymbol{x}_{n-1})] \cdot [\mathrm{tr}\{ \boldsymbol{P}_{n-1|n-1}^{xx} \hat{\boldsymbol{G}}_{n-1|n-1} \} \boldsymbol{f}^{\mathrm{T}}(\boldsymbol{x}_{n-1})] + \boldsymbol{HOT} \tag{7-54}$$

$$\hat{\boldsymbol{F}}_{n-1|n-1} \triangleq [\boldsymbol{\nabla}_{x_{n-1}} \boldsymbol{f}^{\mathrm{T}}(\boldsymbol{x}_{n-1})]^{\mathrm{T}} \Big|_{x_{n-1}=\hat{x}_{n-1|n-1}}$$

$$= \begin{bmatrix} \dfrac{\partial f_1}{\partial x_{n-1,1}} & \cdots & \dfrac{\partial f_1}{\partial x_{n-1,n_x}} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial f_{n_x}}{\partial x_{n-1,1}} & \cdots & \dfrac{\partial f_{n_x}}{\partial x_{n-1,n_x}} \end{bmatrix}_{\boldsymbol{x}_{n-1}=\hat{\boldsymbol{x}}_{n-1|n-1}} \tag{7-55}$$

式中:HOT 代表高阶交叉项干扰。

以同样的方式,观测值预测估计 $\hat{z}_{n|n-1}$、$\boldsymbol{P}^{zz}_{n|n-1}$ 和 $\boldsymbol{P}^{xz}_{n|n-1}$ 变为

$$\hat{\boldsymbol{z}}_{n|n-1} = \boldsymbol{h}(\hat{\boldsymbol{x}}_{n|n-1}) + \frac{1}{2}\mathrm{tr}\{\boldsymbol{P}^{xx}_{n|n-1}\hat{\boldsymbol{H}}_{n|n-1}\}\boldsymbol{h}(\boldsymbol{x}_n)\Big|_{\boldsymbol{x}_n=\hat{\boldsymbol{x}}_{n|n-1}} \tag{7-56}$$

$$\begin{aligned} \boldsymbol{P}^{zz}_{n|n-1} = {} & \boldsymbol{R} + \hat{\boldsymbol{H}}_{n|n-1}\boldsymbol{P}^{xx}_{n|n-1}\hat{\boldsymbol{H}}^{\mathrm{T}}_{n|n-1} + \\ & \frac{1}{2}[\mathrm{tr}\{\boldsymbol{P}^{xx}_{n|n-1}\hat{\boldsymbol{G}}_{n|n-1}\}\boldsymbol{h}(\boldsymbol{x}_n)] \cdot \\ & [\mathrm{tr}\{\boldsymbol{P}^{xx}_{n|n-1}\hat{\boldsymbol{G}}_{n|n-1}\}\boldsymbol{h}^{\mathrm{T}}(\boldsymbol{x}_n)] + \boldsymbol{HOT} \end{aligned} \tag{7-57}$$

$$\boldsymbol{P}^{xz}_{n|n-1} = \boldsymbol{P}^{xx}_{n|n-1}\hat{\boldsymbol{H}}^{\mathrm{T}}_{n|n-1} \tag{7-58}$$

$$\hat{\boldsymbol{H}}_{n|n-1} = [\boldsymbol{\nabla}_{\boldsymbol{x}_n}\boldsymbol{h}^{\mathrm{T}}(\boldsymbol{x}_n)]^{\mathrm{T}}\Big|_{\boldsymbol{x}_n=\hat{\boldsymbol{x}}_{n|n-1}} = \begin{bmatrix} \dfrac{\partial h_1}{\partial x_{n,1}} & \cdots & \dfrac{\partial h_1}{\partial x_{n,n_x}} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial h_{n_x}}{\partial x_{n,1}} & \cdots & \dfrac{\partial h_{n_x}}{\partial x_{n,n_x}} \end{bmatrix}_{\boldsymbol{x}_n=\hat{\boldsymbol{x}}_{n|n-1}} \tag{7-59}$$

$$\begin{aligned} \hat{\boldsymbol{G}}_{n|n-1} & \triangleq \boldsymbol{\nabla}_{\boldsymbol{x}_n}\ \boldsymbol{\nabla}^{\mathrm{T}}_{\boldsymbol{x}_n} \\ & = \begin{bmatrix} \dfrac{\partial^2}{\partial x^2_{n,1}} & \cdots & \dfrac{\partial}{\partial x_{n,1}}\dfrac{\partial}{\partial x_{n,n_x}} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial}{\partial x_{n,n_x}}\dfrac{\partial}{\partial x_{n,1}} & \cdots & \dfrac{\partial^2}{\partial x^2_{n,n_x}} \end{bmatrix}_{\boldsymbol{x}_n=\hat{\boldsymbol{x}}_{n|n-1}} \end{aligned} \tag{7-60}$$

7.2.5 线性化多维扩展卡尔曼滤波器过程

上面推导的 EKF 方程包含比线性更高次的项。标准线性 EKF 可以通过截断,即只保留线性项来得。因此线性化 EKF 和第6章中给出的线性卡尔曼滤波器具有相同形式。

比较表6－1的 LKF 预测方程和表7－2的卡尔曼滤波器预测方程,可以看

出，两者在形式上唯一的区别是，EKF 需要计算雅克比矩阵 $\hat{\boldsymbol{F}}_{n-1|n-1}$ 和 $\hat{\boldsymbol{H}}_{n|n-1}$。在更新步骤，$z_n^{\mathrm{O}}$ 为 t_n 时刻的观测值。

表 7－2　多维扩展卡尔曼滤波器过程

步骤(1)	滤波器初始化	初始化 $\hat{\boldsymbol{x}}_0$ 和 $\boldsymbol{P}_0^{xx}$
步骤(2)	状态矢量预测	$\hat{\boldsymbol{x}}_{n\|n-1}=\boldsymbol{f}(\hat{\boldsymbol{x}}_{n-1\|n-1})$ $\hat{\boldsymbol{F}}_{n-1\|n-1}\triangleq[\boldsymbol{\nabla}_{\boldsymbol{x}_{n-1}}\boldsymbol{f}^{\mathrm{T}}(\boldsymbol{x}_{n-1})]^{\mathrm{T}}_{\boldsymbol{x}_{n-1}=\hat{\boldsymbol{x}}_{n-1\|n-1}}$ $\boldsymbol{P}^{xx}_{n\|n-1}=\hat{\boldsymbol{F}}_{n-1\|n-1}\boldsymbol{P}^{xx}_{n-1\|n-1}\hat{\boldsymbol{F}}^{\mathrm{T}}_{n-1\|n-1}+\boldsymbol{Q}$
步骤(3)	观测值预测	$\hat{\boldsymbol{z}}_{n\|n-1}=\boldsymbol{h}(\hat{\boldsymbol{x}}_{n\|n-1})$ $\hat{\boldsymbol{H}}_{n\|n-1}=[\boldsymbol{\nabla}_{\boldsymbol{x}_n}\boldsymbol{h}^{\mathrm{T}}(\boldsymbol{x}_n)]^{\mathrm{T}}_{\boldsymbol{x}_n=\hat{\boldsymbol{x}}_{n\|n-1}}$ $\boldsymbol{P}^{zz}_{n\|n-1}=\hat{\boldsymbol{H}}_{n\|n-1}\boldsymbol{P}^{xx}_{n\|n-1}\hat{\boldsymbol{H}}^{\mathrm{T}}_{n\|n-1}+\boldsymbol{R}$ $\boldsymbol{P}^{xz}_{n\|n-1}=\boldsymbol{P}^{xx}_{n\|n-1}\hat{\boldsymbol{H}}^{\mathrm{T}}_{n\|n-1}$
步骤(4)	卡尔曼滤波器更新	$\boldsymbol{K}_n\triangleq\boldsymbol{P}^{xz}_{n\|n-1}(\boldsymbol{P}^{zz}_{n\|n-1})^{-1}$ $\hat{\boldsymbol{x}}_{n\|n}=\hat{\boldsymbol{x}}_{n\|n-1}+\boldsymbol{K}_n(z_n^{\mathrm{O}}-\hat{z}_{n\|n-1})$ $\boldsymbol{P}^{xx}_{n\|n}=\boldsymbol{P}^{xx}_{n\|n-1}-\boldsymbol{K}_n\boldsymbol{P}^{zz}_{n\|n-1}\boldsymbol{K}_n^{\mathrm{T}}$
步骤(5)	存储结果	将 $\hat{\boldsymbol{x}}_{n\|n}$ 和 $\boldsymbol{P}^{xx}_{n\|n}$ 存储到跟踪文件中
步骤(6)	返回步骤(2)	

7.2.6　二阶扩展卡尔曼滤波器

有时，强非线性度导致线性化 EKF 无法收敛。在这种情况下，一种方法是使用二阶扩展卡尔曼滤波器(SOEKF)[2,3]。SOEKF 预测方程包括二阶项，因此，用二阶预测方程代替表 7－2 中线性化 EKF 的预测方程，则有

$$\hat{\boldsymbol{x}}_{n|n-1}=\boldsymbol{f}(\hat{\boldsymbol{x}}_{n-1|n-1})+\frac{1}{2}\mathrm{tr}\{\boldsymbol{P}^{xx}_{n-1|n-1}\hat{\boldsymbol{G}}_{n-1|n-1}\}\boldsymbol{f}(\boldsymbol{x}_{n-1})\Big|_{\boldsymbol{x}_{n-1}=\hat{\boldsymbol{x}}_{n-1|n-1}}\tag{7-61}$$

$$\begin{aligned}\boldsymbol{P}^{xx}_{n|n-1}=\boldsymbol{Q}+\hat{\boldsymbol{F}}_{n-1|n-1}\boldsymbol{P}^{xx}_{n-1|n-1}\hat{\boldsymbol{F}}^{\mathrm{T}}_{n-1|n-1}+&\\ \frac{1}{2}[\mathrm{tr}\{\boldsymbol{P}^{xx}_{n-1|n-1}\hat{\boldsymbol{G}}_{n-1|n-1}\}\boldsymbol{f}(\boldsymbol{x}_{n-1})]\cdot&\\ [\mathrm{tr}\{\boldsymbol{P}^{xx}_{n-1|n-1}\hat{\boldsymbol{G}}_{n-1|n-1}\}\boldsymbol{f}^{\mathrm{T}}(\boldsymbol{x}_{n-1})]&\end{aligned}\tag{7-62}$$

$$\hat{\boldsymbol{z}}_{n|n-1}=\boldsymbol{h}(\hat{\boldsymbol{x}}_{n|n-1})+\frac{1}{2}\mathrm{tr}\{\boldsymbol{P}^{xx}_{n|n-1}\hat{\boldsymbol{H}}_{n|n-1}\}\boldsymbol{h}(\boldsymbol{x}_n)\Big|_{\boldsymbol{x}_n=\hat{\boldsymbol{x}}_{n|n-1}}\tag{7-63}$$

$$\begin{aligned}\boldsymbol{P}^{zz}_{n|n-1}=\boldsymbol{R}+\hat{\boldsymbol{H}}_{n|n-1}\boldsymbol{P}^{xx}_{n|n-1}\boldsymbol{H}^{\mathrm{T}}_{n|n-1}+&\\ \frac{1}{2}[\mathrm{tr}\{\boldsymbol{P}^{xx}_{n|n-1}\hat{\boldsymbol{G}}_{n|n-1}\}\boldsymbol{h}(\boldsymbol{x}_n)]\cdot&\end{aligned}$$

$$[\mathrm{tr}\{\boldsymbol{P}_{n|n-1}^{xx}\hat{\boldsymbol{G}}_{n|n-1}\}\boldsymbol{h}^{\mathrm{T}}(\boldsymbol{x}_n)] \tag{7-64}$$

$$\boldsymbol{P}_{n|n-1}^{xz} = \boldsymbol{P}_{n|n-1}^{xx}\hat{\boldsymbol{H}}_{n|n-1}^{\mathrm{T}} \tag{7-65}$$

这些通常很难实现,因为需要计算雅可比矩阵和 Hessian 矩阵。

7.3 多维协方差预测方程的另一种推导

为了以后参考,这里给出 EKF 状态协方差预测方程的另一种推导。在下一章,将用这些推导来求解有限差分,给出有限差分法和无迹卡尔曼滤波器之间的关系。除有限差分方法之外,无迹卡尔曼滤波是用来计算协方差积分的首选数值积分方法。

定义矩阵

$$\tilde{\boldsymbol{g}}(\boldsymbol{c}) \triangleq [\tilde{\boldsymbol{f}}(\boldsymbol{c}) - \hat{\boldsymbol{x}}_{n|n-1}][\tilde{\boldsymbol{f}}(\boldsymbol{c}) - \hat{\boldsymbol{x}}_{n|n-1}]^{\mathrm{T}} \tag{7-66}$$

则式(7－34)可重写为

$$\overline{\boldsymbol{P}}_{n|n-1}^{xx} = \int \tilde{\boldsymbol{g}}(\boldsymbol{c})\mathcal{N}(\boldsymbol{c};\boldsymbol{0},\boldsymbol{I})\mathrm{d}\boldsymbol{c} + \boldsymbol{Q} \tag{7-67}$$

上式有与式(5－51)形式完全相同的状态预测方程,得

$$\overline{\boldsymbol{P}}_{n|n-1}^{xx} = \tilde{\boldsymbol{g}}(\boldsymbol{0}) + \frac{1}{2}\left[\sum_{i=1}^{n_x}\frac{\partial^2}{\partial c_i^2}\tilde{\boldsymbol{g}}(\boldsymbol{c})\right]_{\boldsymbol{c}=\boldsymbol{0}} + \boldsymbol{Q} \tag{7-68}$$

其中 $\overline{\boldsymbol{P}}$ 被用来区分这种形式与式(7－41)。

同样,观测值协方差预测变成

$$\overline{\boldsymbol{P}}_{n|n-1}^{zz} = \tilde{\boldsymbol{p}}(\boldsymbol{0}) + \frac{1}{2}\left[\sum_{i=1}^{n_x}\frac{\partial^2}{\partial c_i^2}\tilde{\boldsymbol{p}}(\boldsymbol{c})\right]_{\boldsymbol{c}=\boldsymbol{0}} + \boldsymbol{R} \tag{7-69}$$

$$\tilde{\boldsymbol{p}}(\boldsymbol{c}) \triangleq [\tilde{\boldsymbol{h}}(\boldsymbol{c}) - \hat{\boldsymbol{z}}_{n|n-1}][\tilde{\boldsymbol{h}}(\boldsymbol{c}) - \hat{\boldsymbol{z}}_{n|n-1}]^{\mathrm{T}} \tag{7-70}$$

互协方差矩阵预测方程(5－41)可以变换成

$$\overline{\boldsymbol{P}}_{n|n-1}^{xz} = \int \tilde{\boldsymbol{r}}(\boldsymbol{c})\mathcal{N}(\boldsymbol{c};\boldsymbol{0},\boldsymbol{I})\mathrm{d}\boldsymbol{c} \tag{7-71}$$

$$\boldsymbol{r}(\boldsymbol{c}) \triangleq [\tilde{\boldsymbol{f}}(\boldsymbol{c}) - \hat{\boldsymbol{x}}_{n|n-1}][\tilde{\boldsymbol{h}}(\boldsymbol{c}) - \hat{\boldsymbol{z}}_{n|n-1}]^{\mathrm{T}} \tag{7-72}$$

则有

$$\overline{\boldsymbol{P}}_{n|n-1}^{xz} = \tilde{\boldsymbol{r}}(\boldsymbol{0}) + \frac{1}{2}\left[\sum_{i=1}^{n_x}\frac{\partial^2}{\partial c_i^2}\tilde{\boldsymbol{r}}(\boldsymbol{c})\right]_{\boldsymbol{c}=\boldsymbol{0}} \tag{7-73}$$

将式(7－66)代入式(7－68),并进行微分,将其结果与表 7－2 中的 $\boldsymbol{P}_{n|n-1}^{xx}$ 进行比较,找出 $\overline{\boldsymbol{P}}_{n|n-1}^{xx}$ 与线性化 EKF 协方差预测估计 $\boldsymbol{P}_{n|n-1}^{xx}$ 之间的关系,即

$$\overline{\boldsymbol{P}}_{n|n-1}^{xx} = \boldsymbol{P}_{n|n-1}^{xx} - \frac{1}{4}\left[\sum_{i=1}^{n_x}\frac{\partial^2}{\partial c_i^2}\tilde{\boldsymbol{f}}(\boldsymbol{c})\right]_{\boldsymbol{c}=\boldsymbol{0}}\left[\sum_{i=1}^{n_x}\frac{\partial^2}{\partial c_i^2}\tilde{\boldsymbol{f}}^{\mathrm{T}}(\boldsymbol{c})\right]_{\boldsymbol{c}=\boldsymbol{0}} \tag{7-74}$$

可以看出，由于$\bar{\boldsymbol{P}}_{n|n-1}^{xx}$包含SOEKF部分高阶项，$\bar{\boldsymbol{P}}_{n|n-1}^{xx}$比线性化EKF得的$\bar{\boldsymbol{P}}_{n|n-1}^{xx}$更为准确。观测值协方差预测估计可以得同样的结果。

在下一章，用这些协方差矩阵形式来推导有限差分滤波器（与将在第9章讨论的无迹卡尔曼滤波器相同），因为本书不会使用 EKF 本身，这里将不再进一步推导。

7.4 EKF 在 DIFAR 船舶跟踪的应用案例分析

利用 EKF 跟踪船舶的笛卡儿位置和速度，当它通过一个浮标场时，建立船舶关于时间的动力学动态模型，以及方位观测值与船舶位置之间的关系模型，还需要讨论两个问题：如何初始化（$\hat{\boldsymbol{x}}_0, \boldsymbol{P}_{n|n-1}^{xx}$）以及如何选择一个合适的加速度噪声 q。

7.4.1 船舶运动动态模型

矢量矩阵形式的动态模型可以由式（4－1）定义的船舶状态矢量得，将式（4－2）、式（4－3）代入下式

$$\boldsymbol{x}_n = \boldsymbol{F}\boldsymbol{x}_{n-1} + \boldsymbol{v}_{n-1} \tag{7-75}$$

这里增加了高斯白噪声加速度噪声项 $\boldsymbol{v}_{n-1}$，以允许船舶跟踪过程中微小的变化和机动操纵。将时间不变转移矩阵 $\boldsymbol{F}$ 定义为

$$\boldsymbol{F} \triangleq \begin{bmatrix} 1 & T & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & T \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{7-76}$$

并给出

$$\boldsymbol{Q} \triangleq \varepsilon\{\boldsymbol{v}_n \boldsymbol{v}_n^{\mathrm{T}}\} = q\begin{bmatrix} \frac{T^3}{3} & \frac{T^2}{2} & 0 & 0 \\ \frac{T^2}{2} & T & 0 & 0 \\ 0 & 0 & \frac{T^3}{3} & \frac{T^2}{2} \\ 0 & 0 & \frac{T^2}{2} & T \end{bmatrix} \tag{7-77}$$

7.4.2 DIFAR 浮标观测模型

定义 t_n 时刻的浮标场观测矢量为

$$z_n = [\theta_{1,n}, \cdots, \theta_{M,n}]^{\mathrm{T}} \tag{7-78}$$

每个浮标的观测模型由式（4－7）定义。然后用如下矢量观测方程联合所有浮

标，有

$$z_n = h(x_n) + w_n \tag{7-79}$$

$$\theta_{m,n} = h_m(x_n) + w_{m,n} \tag{7-80}$$

$$h_m(x_n) = \tan^{-1}\left(\frac{r_n^x - x_m}{r_n^y - y_m}\right) \tag{7-81}$$

高斯白噪声过程观测值的协方差 w_n 定义为

$$R \triangleq \varepsilon\{w_n w_n^{\mathrm{T}}\} = \begin{bmatrix} \sigma_1^2 & 0 & \cdots & 0 & 0 \\ 0 & \sigma_2^2 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \sigma_{M-1}^2 & 0 \\ 0 & 0 & \cdots & 0 & \sigma_M^2 \end{bmatrix} \tag{7-82}$$

式中：σ_m^2 为与第 m 个浮标观测值相关的方差。在模拟仿真中，假定 $\sigma_m = \sigma = 3°, \forall m$。

因为动态方程是线性的，表6－1中的LKF状态预测方程可以用于跟踪滤波器。这里用到的状态预测方程来自于表6－1，即

$$\hat{x}_{n|n-1} = F\hat{x}_{n-1|n-1} \tag{7-83}$$

$$P_{n|n-1}^{xx} = FP_{n-1|n-1}^{xx}F^{\mathrm{T}} + Q \tag{7-84}$$

然而，由于每个浮标的观测方程式(7－81)明显为非线性，必须使用表7－2中的EKF预测方程，而这需要计算雅可比矩阵 $\hat{H}_{n|n-1}$。

雅可比矩阵定义为

$$\hat{H}_{n|n-1} = [\nabla_{x_n} h^{\mathrm{T}}(x_n)]^{\mathrm{T}}\Big|_{x_n = \hat{x}_{n|n-1}}$$

$$= \left[\begin{bmatrix} \dfrac{\partial}{\partial r_n^x} \\ \dfrac{\partial}{\partial v_n^x} \\ \dfrac{\partial}{\partial r_n^y} \\ \dfrac{\partial}{\partial v_n^y} \end{bmatrix} [h_1(x_n) \cdots h_M(x_n)]\right]^{\mathrm{T}}_{x_n = \hat{x}_{n|n-1}}$$

$$= \begin{bmatrix} \dfrac{\partial h_1}{\partial r_n^x} & \cdots & \dfrac{\partial h_1}{\partial v_n^y} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial h_M}{\partial r_n^x} & \cdots & \dfrac{\partial h_M}{\partial v_n^y} \end{bmatrix}_{x_n = \hat{x}_{n|n-1}} \tag{7-85}$$

展开导数项，得

$$\hat{\boldsymbol{H}}_{n|n-1} = \begin{bmatrix} \dfrac{\hat{r}^y_{n|n-1} - y_1}{R^2_{n|n-1}} & \cdots & \dfrac{\hat{r}^y_{n|n-1} - y_M}{R^2_{n|n-1}} \\ 0 & \cdots & 0 \\ \dfrac{\hat{r}^x_{n|n-1} - x_1}{R^2_{n|n-1}} & \cdots & \dfrac{\hat{r}^x_{n|n-1} - x_M}{R^2_{n|n-1}} \\ 0 & \cdots & 0 \end{bmatrix} \tag{7-86}$$

$$R_{n|n-1} = \sqrt{(\hat{r}^x_{n|n-1})^2 + (\hat{r}^y_{n|n-1})^2} \tag{7-87}$$

将 EKF 应用到 DIFAR 案例分析时，所使用的观测预测方程（来自于表7－2）为

$$\hat{z}_{n|n-1} = \boldsymbol{h}(\hat{\boldsymbol{x}}_{n|n-1}) \tag{7-88}$$

$$\boldsymbol{P}^{zz}_{n|n-1} = \hat{\boldsymbol{H}}_{n|n-1}\boldsymbol{P}^{xx}_{n|n-1}\hat{\boldsymbol{H}}^{\mathrm{T}}_{n|n-1} + \boldsymbol{R} \tag{7-89}$$

$$\boldsymbol{P}^{xz}_{n|n-1} = \boldsymbol{P}^{xx}_{n|n-1}\hat{\boldsymbol{H}}^{\mathrm{T}}_{n|n-1} \tag{7-90}$$

当然，这里使用表7－2步骤(4)中标准的卡尔曼滤波器更新方程。

7.4.3 初始化卡尔曼滤波器族中所有滤波器

在后面章节，将对卡尔曼滤波器类中所有跟踪滤波器性能进行比较。为确保比较的合理性，必须对所有的跟踪滤波器进行同样的初始化。首先，假定目标船舶由西南朝着原点径向入站，如图7－1所示，最初的距离 R 为 4n mile。

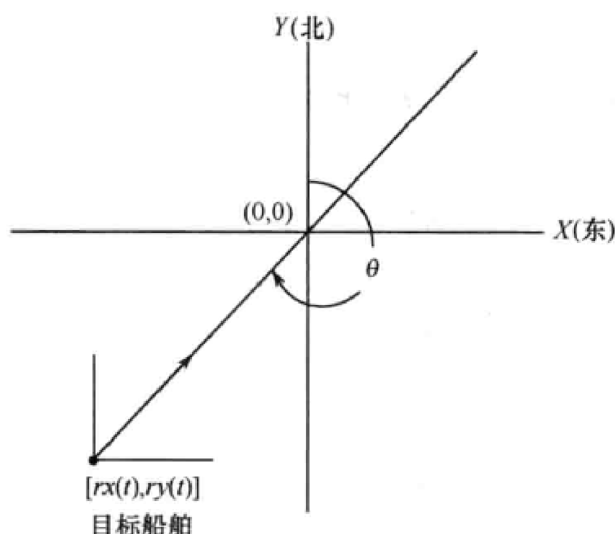


图7－1　初始化几何示意图

给定一组初始浮标测量值$\{\theta_m, m = 1,\cdots,M\}$，用下式初始化状态矢量位置分量，即

$$r^x_0 = \mathrm{mean}(r^x_m - R\sin\theta_m) \tag{7-91}$$

$$r^y_0 = \mathrm{mean}(r^y_m - R\cos\theta_m) \tag{7-92}$$

式中：(r^x_m, r^y_m)为第 m 个浮标的位置；mean 为所有浮标的平均值。

根据速度分量，对所有浮标的第一组方位测量值求平均，生成目标相对于原点的方位估计 $\theta_0 = [\sum \theta_m]/M$。任意设定目标船舶的速度 $v = 30\text{kn}$，用 $\theta_0 \simeq \vartheta - \pi$ 计算目标航向，这里 v 为目标航向。

初始目标速度分量可以由下式计算，即

$$v_0^x = v \sin\vartheta \tag{7-93}$$

$$v_0^y = v \cos \vartheta \tag{7-94}$$

基于初始化过程，可以发现，如果想由第一组 DIFAR 元素方位测量值得初始方位估计，必须确定两个变量，即距离 R 和速度 v。

因为没有足够的信息由测量值协方差生成初始协方差矩阵，任意指定一个对角矩阵，用来初始化状态误差协方差，即

$$\boldsymbol{P}_0^{xx} = \begin{bmatrix} \sigma_x^2 & 0 & 0 & 0 \\ 0 & \sigma_{vx}^2 & 0 & 0 \\ 0 & 0 & \sigma_y^2 & 0 \\ 0 & 0 & 0 & \sigma_{vy}^2 \end{bmatrix} \tag{7-95}$$

式中：$\sigma_x{}^2 = \sigma_y{}^2 = 0.2$；$\sigma_{vx}^2 = \sigma_{vy}^2 = 3$。

7.4.4 选定加速度噪声参数

如前所述，q 的单位为 v^2/t 或 $a^2 \times t$。对于像船舶这样运动缓慢的目标，使用速度的变化值比较方便；而对于像战斗机或导弹这种快速运动的目标，使用加速度的变化值（如3g 转向，g 为重力加速度）更为适合。对于缓慢运动的船舶目标，选择 $q = 10^{-6}$ 以允许船舶速度由于波浪作用和风力引起的微小偏差。q 取值很小，是假定在跟踪间隔中，没有机动操纵船舶得的结果。对于一个机动操纵的船舶，取较高的 q 值更为合适。

7.4.5 EKF 跟踪滤波器结果

运行 EKF 跟踪滤波器，设滤波更新率为 1s、船舶接收信号集合的信噪比为 20dB、时间带宽积为90，得如图7－2 所示的跟踪结果。这个航迹图直观地给出了跟踪算法的执行情况，后续章节将给出深入的性能测量讨论与对比。如图所示，由于初始化不够准确，跟踪器需要一段时间才能收敛于真实航迹，但收敛于真实航迹（用灰色表示）的速度很快。如果初始化不够精确，跟踪器的容错就非常重要。因为滤波器能够利用观测值迅速地纠正初始化错误，EKF（以及后面所有的非线性卡尔曼滤波器）与初始值 R 和 v 的选择几乎无关。运行滤波器，将信噪比设为20dB，距离初始值为20 ～ 3n mile，速度初始化为10 ～ 30kn，滤波器迅速收敛于真实航迹。

在第 4 章中给出，当信噪比很低时，观测值的方位噪声变成非高斯分布的。

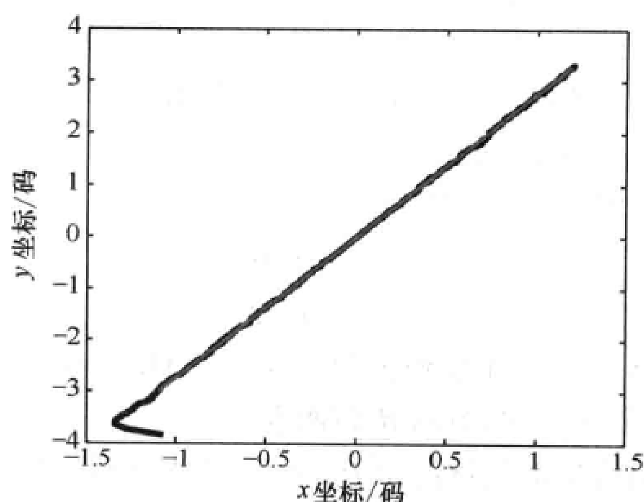


图7－2　通过浮标场船舶的估计航迹与真实航迹对比

图7－3表明输入信号信噪比对高斯EKF航迹估计的影响，信噪比从－5到20dB，增量为5dB。在图7－3中，黑色线条是100次蒙特卡洛运行（采用同样的初始值）得的滤波器跟踪估计，灰色线条为真实航迹。第14章中将讨论，在一组蒙特卡洛运行中发散航迹比例的性能度量。在图7－3中列出每种信噪比下发散航迹的数目，由于有100次蒙特卡洛运行，因此发散航迹的数目与发散航迹的百分比相同。从图中很容易看到，当信噪比小于等于5dB时，由于非高斯噪声导致的方位观测异常值过多，EKF跟踪非常困难。

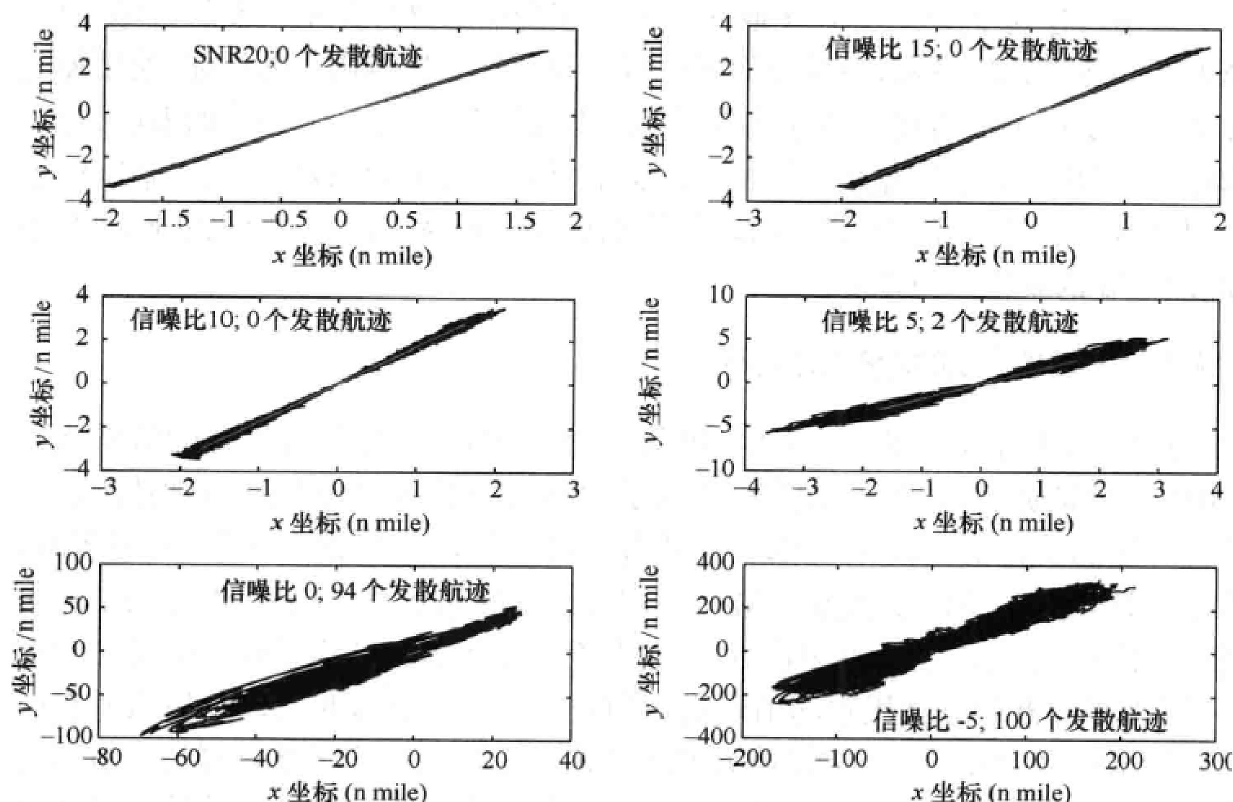


图7－3　6种信噪比下EKF跟踪器输出结果比较

参考文献

[1] Jazwinski AH. *Stochastic Processes and Filtering Theory.* Academic Press, 1970.

[2] Gelb A. *Applied Optimal Estimation.* The MIT Press, 1974.

[3] Mahakababis A, Farooq M. A Second - order method for state estimation of non - linear dynamical systems. *Int. J. Control*, 1971, 14(4): 631 - 639.

第8章
Sigma 点类别:有限差分卡尔曼滤波器

如果非线性程度不是很严重,线性化 EKF 可以得合理的估计结果。对于动态或观测转移函数为高度非线性的情况,可以在 EKF 中包含二阶项,来实现估计。由于需要额外计算 Hessian 矩阵,这种增加的复杂性有时很难实现。解决这个问题的一种方法是用等价的有限差分取代 EKF 中的微分运算。

Schei 首次提出,在非线性估计算法中,利用中心有限差分法实现线性化[1,2]。EKF 为函数 $\boldsymbol{f}(\boldsymbol{x})$ 在点 $\hat{\boldsymbol{x}}$ 处的线性化,则 Schei 方法为在中心差分点 $\hat{\boldsymbol{x}}+\boldsymbol{D}\boldsymbol{q}$ 处的线性化。由于 Schei 只取代了状态(观测)矢量及其协方差预测方程中的雅可比矩阵,所以这种方法仅在状态(观测)矢量预测步骤方面比 EKF 更为精确。

Norgaard 等提出一种更为精确且无需求导的非线性系统估计方法[3,4],用多维 Stirling 插值公式代替非线性函数的泰勒多项式逼近,这样能够在状态(观测)矢量和协方差预测步骤上保持高阶项,本章内容在文献[3]基础上提出。

Ito、Xiong[5] 与 Wu 等[6] 的论文推广了有限差分逼近方法,将其纳入预测技术框架,采用更一般的多项式展开方法,来实现非线性函数的逼近。

本章给出有限差分卡尔曼滤波器的完整推导,从标量情况开始,然后推广至多维情况。

8.1 一维有限差分卡尔曼滤波器

8.1.1 一维有限差分状态预测

式(7-10)给出了 $\tilde{f}(c)$ 的 EKF 状态预测方程标量推导,为清晰起见,这里再重复一次。

$$\hat{x}_{n|n-1} = \tilde{f}(0) + \frac{1}{2}\left[\frac{\mathrm{d}^2}{\mathrm{d}c^2}\tilde{f}(c)\right]_{c=0} \tag{8-1}$$

如前所述,c 具有三阶精度。

用式(2－47)有限差分近似法替换二阶导数,则(8－1)变为

$$\hat{x}_{n|n-1} = \tilde{f}(0) + \frac{1}{2q^2}[\tilde{f}(q) - 2\tilde{f}(0) + \tilde{f}(-q)] \tag{8-2}$$

合并相关项,式(8－2)可重写为

$$\hat{x}_{n|n-1} = \left[\frac{q^2-1}{q^2}\right]\tilde{f}(0) + \frac{1}{2q^2}[\tilde{f}(q) + \tilde{f}(-q)] \tag{8-3}$$

令

$$w_0 \triangleq \frac{q^2-1}{q^2} \tag{8-4}$$

则有

$$q = \frac{1}{\sqrt{1-w_0}} \tag{8-5}$$

有限差分计算点定义为

$$c^{(j)} = qr^{(j)} \tag{8-6}$$

$$r^{(j)} \triangleq \begin{cases} 0 & j=0 \\ [1] \in \mathbf{R}^1 & j=1,2 \end{cases} \tag{8-7}$$

$[1] \in \mathbf{R}^1$ 符号的解释可参阅第2.2节。利用式(8－4)和式(8－6),式(8－3)变为

$$\hat{x}_{n|n-1} = w_0\tilde{f}(c^{(0)}) + \frac{1-w_0}{2}[\tilde{f}(c^{(1)}) + \tilde{f}(c^{(2)})] \tag{8-8}$$

用离散值取代式(7－3)中定义的连续 c 值,得

$$\tilde{f}(c^{(j)}) \triangleq f(\hat{x}_{n-1|n-1} + \sigma_{xx,n-1|n-1}c^{(j)}) \tag{8-9}$$

定义一维有限差分 Sigma 点,得有限差分计算点的$\chi_{n-1|n-1}^{(j)}$ 项,即

$$\chi_{n-1|n-1}^{(j)} \triangleq \hat{x}_{n-1|n-1} + \sigma_{xx,n-1|n-1}c^{(j)}$$
$$= \begin{cases} \hat{x}_{n-1|n-1} & j=0 \\ \hat{x}_{n-1|n-1} + \dfrac{1}{\sqrt{1-w_0}}\sigma_{xx,n-1|n-1} & j=1 \\ \hat{x}_{n-1|n-1} - \dfrac{1}{\sqrt{1-w_0}}\sigma_{xx,n-1|n-1} & j=2 \end{cases} \tag{8-10}$$

这样,式(8－8)一维有限差分状态预测方程变为

$$\hat{x}_{n|n-1} = w_0 f(\chi_{n-1|n-1}^{(0)}) + \frac{1-w_0}{2}[f(\chi_{n-1|n-1}^{(1)}) + f(\chi_{n-1|n-1}^{(2)})] \tag{8-11}$$

令 $\omega_1 = \omega_2 \triangleq (1-\omega_0)/2$,式(8－11)可以写成 Sigma 点的和,即

$$\hat{x}_{n|n-1} = \sum_{j=0}^{2} w_j f(\chi_{n-1|n-1}^{(j)}) \tag{8-12}$$

8.1.2 一维有限差分状态预测

用式(2－44)和式(2－47)中的有限差分等效代替式(7－17)的一阶和二阶导数,状态误差方差预测方程变为

$$\sigma_{xx,n|n-1}^2 = \frac{1}{4q^2}[\tilde{f}(q) - \tilde{f}(-q)]^2 + \frac{1}{2q^4}[\tilde{f}(q) - 2\tilde{f}(0) + \tilde{f}(-q)]^2 + \sigma_{v,n}^2 \tag{8-13}$$

利用式(8－4)、式(8－6)和式(8－10),式(8－13)可简化为

$$\begin{aligned}\sigma_{xx,n|n-1}^2 = &\frac{1-w_0}{4}[f(\chi_{n-1|n-1}^{(1)}) - f(\chi_{n-1|n-1}^{(2)})]^2 + \\ &\frac{(1-w_0)^2}{2}[f(\chi_{n-1|n-1}^{(1)}) - 2f(\chi_{n-1|n-1}^{(0)}) + \\ &+ f(\chi_{n-1|n-1}^{(2)})]^2 + \sigma_{v,n}^2\end{aligned} \tag{8-14}$$

注意,方程具有5阶精度,这在第7章已说明。

8.1.3 一维有限差分观测预测方程

采用与前面两节同样的步骤,给出一维观测预测和互方差预测方程为

$$\hat{z}_{n|n-1} = \sum_{j=0}^{2} w_i h\ (\chi_{n|n-1}^{(j)}) \tag{8-15}$$

$$\begin{aligned}\sigma_{zz,n|n-1}^2 = &\frac{1-w_0}{4}[h(\chi_{n|n-1}^{(1)}) - h(\chi_{n|n-1}^{(2)})]^2 + \\ &\frac{(1-w_0)^2}{2}[h(\chi_{n|n-1}^{(1)}) - 2h(\chi_{n|n-1}^{(0)}) + h(\chi_{n|n-1}^{(2)})]^2 + \sigma_{w,n}^2\end{aligned} \tag{8-16}$$

$$\sigma_{xz,n|n-1}^2 = \sigma_{xx,n|n-1}^2\{w_1[h(\chi_{n|n-1}^{(1)}) - h(\chi_{n|n-1}^{(2)})]\} \tag{8-17}$$

$$\chi_{n|n-1}^{(j)} = \begin{cases}\hat{x}_{n|n-1} & j=0 \\ \hat{x}_{n|n-1} + \dfrac{1}{\sqrt{1-w_0}}\sigma_{xx,n|n-1}^2 & j=1 \\ \hat{x}_{n|n-1} - \dfrac{1}{\sqrt{1-w_0}}\sigma_{xx,n|n-1}^2 & j=2\end{cases} \tag{8-18}$$

8.1.4 一维有限差分卡尔曼滤波器过程

式(8－12)、式(8－14)～式(8－17)构成了一维有限差分卡尔曼滤波器的预测部分,其状态和观测预测具有三次精度,协方差预测具有五次精度。表8－1给出了完整的一维有限差分卡尔曼滤波器过程。

8.1.5 简化的一维有限差分预测方程

对于一维预测方程可以进行简化，最简单的方法是取 $\omega_0 = 0$，这样可以去掉位于原点的点，得没有自由参数的预测方程，即

表 8－1　一维有限差分卡尔曼滤波器过程

步骤(1)	滤波器初始化	设置参数 $0 \leqslant \omega_0 < 1$ 初始化和 $\hat{x}_0$ 和 $\sigma^2_{xx,0}$
步骤(2)	状态矢量预测	$\chi^{(j)}_{n-1\mid n-1} = \begin{cases} \hat{x}_{n-1\mid n-1} & j=0 \\ \hat{x}_{n-1\mid n-1} + \dfrac{1}{\sqrt{1-w_0}}\sigma_{xx,n-1\mid n-1} & j=1 \\ \hat{x}_{n-1\mid n-1} - \dfrac{1}{\sqrt{1-w_0}}\sigma_{xx,n-1\mid n-1} & j=2 \end{cases}$ $\hat{x}_{n\mid n-1} = \sum_{j=0}^{2} w_j f(\chi^{(j)}_{n-1\mid n-1})$ $\sigma^2_{xx,n\mid n-1} = \dfrac{1-w_0}{4} \cdot$ $[f(\chi^{(1)}_{n-1\mid n-1}) - f(\chi^{(2)}_{n-1\mid n-1})]^2 +$ $\dfrac{(1-w_0)^2}{2}[f(\chi^{(1)}_{n-1\mid n-1}) - 2f(\chi^{(0)}_{n-1\mid n-1}) +$ $f(\chi^{(2)}_{n-1\mid n-1})]^2 + \sigma^2_{v,n}$
步骤(3)	观测值预测	$\chi^{(j)}_{n\mid n-1} = \begin{cases} \hat{x}_{n\mid n-1} & j=0 \\ \hat{x}_{n\mid n-1} + \dfrac{1}{\sqrt{1-w_0}}\sigma^2_{xx,n\mid n-1} & j=1 \\ \hat{x}_{n\mid n-1} - \dfrac{1}{\sqrt{1-w_0}}\sigma^2_{xx,n\mid n-1} & j=2 \end{cases}$ $\hat{z}_{n\mid n-1} = \sum_{j=0}^{2} w_i h(\chi^{(j)}_{n\mid n-1})$ $\sigma^2_{zz,n\mid n-1} = \dfrac{1-w_0}{4} \cdot$ $[h(\chi^{(1)}_{n\mid n-1}) - h(\chi^{(2)}_{n\mid n-1})]^2 + \dfrac{(1-w_0)^2}{2} \cdot$ $[h(\chi^{(1)}_{n\mid n-1}) - 2h(\chi^{(0)}_{n\mid n-1}) + h(\chi^{(2)}_{n\mid n-1})]^2 + \sigma^2_{w,n}$ $\sigma^2_{xz,n\mid n-1} = \sigma^2_{xx,n\mid n-1} \cdot$ $\{w_1[h(\chi^{(1)}_{n\mid n-1}) - h(\chi^{(2)}_{n\mid n-1})]\}$
步骤(4)	卡尔曼滤波器更新	$K_n = \sigma^2_{xz,n\mid n-1} / \sigma^2_{zz,n\mid n-1}$ $\hat{x}_{n\mid n} = \hat{x}_{n\mid n-1} + K_n(z^o_n - \hat{z}_{n\mid n-1})$ $\sigma^2_{xx,n\mid n} = \sigma^2_{xx,n\mid n-1} - K^2_n \sigma^2_{zz,n\mid n-1}$
步骤(5)	存储结果	将 $\hat{\boldsymbol{x}}_{n\mid n}$ 和 $\boldsymbol{P}^{xx}_{n\mid n}$ 和存储到跟踪文件中
步骤(6)	返回步骤(2)	

$$\hat{x}_{n|n-1} = \frac{1}{2}[f(\chi_{n-1|n-1}^{(1)}) + f(\chi_{n-1|n-1}^{(2)})] \tag{8-19}$$

$$\begin{aligned}\sigma_{xx,n|n-1}^2 = &\frac{1}{4}[f(\chi_{n-1|n-1}^{(1)}) - f(\chi_{n-1|n-1}^{(2)})]^2 + \\ &\frac{1}{2}[f(\chi_{n-1|n-1}^{(1)}) - 2f(\chi_{n-1|n-1}^{(0)}) + \\ &f(\chi_{n-1|n-1}^{(2)})]^2 + \sigma_{v,n}^2\end{aligned} \tag{8-20}$$

$$\hat{z}_{n|n-1} = \frac{1}{2}[h(\chi_{n|n-1}^{(1)}) + h(\chi_{n|n-1}^{(2)})] \tag{8-21}$$

$$\begin{aligned}\sigma_{zz,n|n-1}^2 = &\frac{1}{4}[h(\chi_{n|n-1}^{(1)}) - h(\chi_{n|n-1}^{(2)})]^2 + \\ &\frac{1}{2}[h(\chi_{n|n-1}^{(1)}) - 2h(\chi_{n|n-1}^{(0)}) + \\ &h(\chi_{n|n-1}^{(2)})]^2 + \sigma_{w,n}^2\end{aligned} \tag{8-22}$$

$$\sigma_{xz,n|n-1}^2 = \frac{1}{2}\sigma_{xx,n|n-1}[h(\chi_{n|n-1}^{(1)}) - h(\chi_{n|n-1}^{(2)})] \tag{8-23}$$

$$\chi_{n-1|n-1}^{(j)} = \begin{cases}\hat{x}_{n-1|n-1} + \sigma_{xx,n-1|n-1} & j = 1\\ \hat{x}_{n-1|n-1} - \sigma_{xx,n-1|n-1} & j = 2\end{cases} \tag{8-24}$$

$$\chi_{n|n-1}^{(j)} = \begin{cases}\hat{x}_{n|n-1} + \sigma_{xx,n|n-1} & j = 1\\ \hat{x}_{n|n-1} - \sigma_{xx,n|n-1} & j = 2\end{cases} \tag{8-25}$$

另一种简化方法为截断高阶项。在下一节处理多维情况时,将更详细地讨论这种方法。

8.2 多维有限差分卡尔曼滤波器

8.2.1 多维有限差分状态预测

与一维情况采取相似的方式,首先重复一下式(7 - 33)给出的 EKF 多维状态预测方程,即

$$\hat{\boldsymbol{x}}_{n|n-1} = \tilde{\boldsymbol{f}}(\boldsymbol{0}) + \sum_{j=1}^{n_x}\left[\frac{\partial^2}{\partial c_j^2}\tilde{\boldsymbol{f}}(\boldsymbol{c})\right]_{\boldsymbol{c}=\boldsymbol{0}} \tag{8-26}$$

使用式(2 - 74)给出的多维 Stirling 多项式的二阶有限差分项,令 $\boldsymbol{x} \to \boldsymbol{c}$ 且 $\boldsymbol{x}_0 = \boldsymbol{0}$,则式(8 - 26)变为

$$\begin{aligned}\hat{\boldsymbol{x}}_{n|n-1} &= \tilde{\boldsymbol{f}}(\boldsymbol{0}) + \frac{1}{2q^2}\sum_{j=1}^{n_x}[\tilde{\boldsymbol{f}}(q\boldsymbol{e}_j) - 2\tilde{\boldsymbol{f}}(\boldsymbol{0}) + \tilde{\boldsymbol{f}}(-q\boldsymbol{e}_j)]\\ &= \left[\frac{q^2 - n_x}{q^2}\right]\tilde{\boldsymbol{f}}(\boldsymbol{0}) + \frac{1}{2q^2}\sum_{j=1}^{n_x}[\tilde{\boldsymbol{f}}(q\boldsymbol{e}_j) + \tilde{\boldsymbol{f}}(-q\boldsymbol{e}_j)]\end{aligned} \tag{8-27}$$

式中：$\boldsymbol{e}_j$ 为第 j 维笛卡儿坐标轴方向的单位矢量；q 为步长，要求必须为大于零的有限实数值。

定义

$$w_0 \triangleq \frac{q^2 - n_x}{q^2} \tag{8-28}$$

得等式

$$\frac{1}{q^2} = \frac{1 - w_0}{n_x} \tag{8-29}$$

$$q = \sqrt{\frac{n_x}{1 - w_0}} \tag{8-30}$$

这里 ω_0 为一自由参数，决定了 q 的值。为保证 q 为有限大于零的实数，必须限制 ω_0 的取值范围 $0 \leqslant \omega_0 \leqslant 1$。

这样，式(8－27) 可写为

$$\hat{\boldsymbol{x}}_{n|n-1} = w_0 \tilde{\boldsymbol{f}}(\boldsymbol{0}) + \frac{1 - w_0}{2n_x} \sum_{j=1}^{n_x} \left[\tilde{\boldsymbol{f}}\left(\sqrt{\frac{n_x}{1 - w_0}} \boldsymbol{e}_j \right) + \tilde{\boldsymbol{f}}\left(- \sqrt{\frac{n_x}{1 - w_0}} \boldsymbol{e}_j \right) \right] \tag{8-31}$$

为进一步简化，利用 2.2 节定义的多维矢量生成器函数[7,8]。例如，在四维的情况下，有

$$\boldsymbol{r} = \begin{cases} [0] \in \mathbf{R}^4 \\ [1] \in \mathbf{R}^4 \end{cases} = \begin{cases} \boldsymbol{r}^{(0)} = [0,0,0,0]^{\mathrm{T}} \\ \boldsymbol{r}^{(1)} = [1,0,0,0]^{\mathrm{T}} = \boldsymbol{e}_1 \\ \boldsymbol{r}^{(2)} = [0,1,0,0]^{\mathrm{T}} = \boldsymbol{e}_2 \\ \boldsymbol{r}^{(3)} = [0,0,1,0]^{\mathrm{T}} = \boldsymbol{e}_3 \\ \boldsymbol{r}^{(4)} = [0,0,0,1]^{\mathrm{T}} = \boldsymbol{e}_4 \\ \boldsymbol{r}^{(5)} = [-1,0,0,0]^{\mathrm{T}} = -\boldsymbol{e}_1 \\ \boldsymbol{r}^{(6)} = [0,-1,0,0]^{\mathrm{T}} = -\boldsymbol{e}_2 \\ \boldsymbol{r}^{(7)} = [0,0,-1,0]^{\mathrm{T}} = -\boldsymbol{e}_3 \\ \boldsymbol{r}^{(8)} = [0,0,0,-1]^{\mathrm{T}} = -\boldsymbol{e}_4 \end{cases} \tag{8-32}$$

这些矢量积分点形成一个四维网格，包含沿笛卡儿坐标系坐标轴的单位矢量和在一个在原点的点。如果排除位于原点的点，这些单位矢量都落在半径为 1 的四维超球体 $\mathcal{N}(\boldsymbol{c};\boldsymbol{0},\boldsymbol{I})$ 的外轮廓上。二维图形化示例如图 2－9 所示。

对于一般的 n_x 维情况，有

$$\boldsymbol{r} = \begin{cases} [0] \in \mathbf{R}^{n_x} & j = 0 \\ [1] \in \mathbf{R}^{n_x} & j = 1, \cdots, 2n_x \end{cases} \tag{8-33}$$

式(2－30) 给出 $[1] \in \mathbf{R}^{n_x}$ 中的矢量点值。离散的 $\boldsymbol{c}$ 定义为

$$\boldsymbol{c}^{(j)} = q\boldsymbol{r}^{(j)} = \sqrt{\frac{n_x}{1 - w_0}} \boldsymbol{r}^{(j)} \quad j = 0, \cdots, 2n_x \tag{8-34}$$

则式(5 - 53)可以写为

$$\tilde{\boldsymbol{f}}(\boldsymbol{c}^{(j)}) = \boldsymbol{f}(\hat{\boldsymbol{x}}_{n-1|n-1} + \boldsymbol{D}_{n-1|n-1}\boldsymbol{c}^{(j)}) \tag{8-35}$$

这里$\boldsymbol{D}_{n-1|n-1}$由式(5 -48)定义。由此得一般的多维有限差分Sigma点的定义,即

$$\begin{aligned}\boldsymbol{\chi}_{n-1|n-1}^{(j)} &\triangleq \hat{\boldsymbol{x}}_{n-1|n-1} + \boldsymbol{D}_{n-1|n-1}\boldsymbol{c}^{(j)} \\ &= \hat{\boldsymbol{x}}_{n-1|n-1} + \sqrt{\frac{n_x}{1-w_0}}\boldsymbol{D}_{n-1|n-1}\boldsymbol{r}^{(j)}\ j = 0,\cdots,2n_x\end{aligned} \tag{8-36}$$

因此,式(8 - 31)可以重写为

$$\hat{\boldsymbol{x}}_{n|n-1} = \sum_{j=0}^{2n_x} w_j \boldsymbol{f}(\chi_{n-1|n-1}^{(j)}) \tag{8-37}$$

$$w_j = \begin{cases} w_0 & j = 0 \\ \dfrac{1-w_0}{2n_x} & j = 1,\cdots,2n_x \end{cases} \tag{8-38}$$

这里$\sum_{j=0}^{2n_x}\omega_j = 1$。如前所述,这是一个三阶逼近。

8.2.2 多维有限差分状态协方差预测

式(7 -41)给出了多维EKF预测状态误差协方差矩阵,为清晰起见,这里再重复一次。

$$\begin{aligned}\boldsymbol{P}_{n|n-1}^{xx} = &\left[\sum_{i=1}^{n_x}\frac{\partial}{\partial c_i}\tilde{\boldsymbol{f}}(\boldsymbol{c})\right]_{\boldsymbol{c}=\boldsymbol{0}}\left[\sum_{i=1}^{n_x}\frac{\partial}{\partial c_i}\tilde{\boldsymbol{f}}^{\mathrm{T}}(\boldsymbol{c})\right]_{\boldsymbol{c}=\boldsymbol{0}} + \\ &\frac{1}{2}\left[\sum_{i=1}^{n_x}\frac{\partial^2}{\partial c_i^2}\tilde{\boldsymbol{f}}(\boldsymbol{c})\right]_{\boldsymbol{c}=\boldsymbol{0}}\left[\sum_{i=1}^{n_x}\frac{\partial^2}{\partial c_i^2}\tilde{\boldsymbol{f}}^{\mathrm{T}}(\boldsymbol{c})\right]_{\boldsymbol{c}=\boldsymbol{0}} + \\ &\left[\sum_{i=1}^{n_x}\sum_{\substack{j=1\\j\neq i}}^{n_x}\frac{\partial^2}{\partial c_i\partial c_j}\tilde{\boldsymbol{f}}(\boldsymbol{c})\right]_{\boldsymbol{c}=\boldsymbol{0}}\left[\sum_{i=1}^{n_x}\sum_{\substack{j=1\\j\neq i}}^{n_x}\frac{\partial^2}{\partial c_i\partial c_j}\tilde{\boldsymbol{f}}^{\mathrm{T}}(\boldsymbol{c})\right]_{\boldsymbol{c}=\boldsymbol{0}}\end{aligned} \tag{8-39}$$

对于$\tilde{\boldsymbol{f}}(\boldsymbol{c})$的每一维,利用方程(2 - 73)和式(2 - 74)给出的多维Stirling多项式的一阶和二阶有限差分项,式(8 - 39)变为

$$\begin{aligned}\boldsymbol{P}_{n|n-1}^{xx} = &\frac{1}{4q^2}\sum_{i=1}^{n_x}\sum_{j=1}^{n_x}[\tilde{\boldsymbol{f}}(q\boldsymbol{e}_i) - \tilde{\boldsymbol{f}}(-q\boldsymbol{e}_i)][\tilde{\boldsymbol{f}}(q\boldsymbol{e}_j) - \tilde{\boldsymbol{f}}(-q\boldsymbol{e}_j)]^{\mathrm{T}} + \\ &\frac{1}{2q^4}\sum_{i=1}^{n_x}\sum_{j=1}^{n_x}[\tilde{\boldsymbol{f}}(q\boldsymbol{e}_i) - 2\tilde{\boldsymbol{f}}(\boldsymbol{0}) + \tilde{\boldsymbol{f}}(-q\boldsymbol{e}_i)]\cdot \\ &[\tilde{\boldsymbol{f}}(q\boldsymbol{e}_j) - 2\tilde{\boldsymbol{f}}(\boldsymbol{0}) + \tilde{\boldsymbol{f}}(-q\boldsymbol{e}_j)]^{\mathrm{T}} + \\ &\frac{1}{4q^4}\sum_{i=1}^{n_x}\sum_{j=1}^{n_x}\sum_{k=1}^{n_x}\sum_{l=1}^{n_x}[\tilde{\boldsymbol{f}}(q\boldsymbol{e}_i + q\boldsymbol{e}_j) - \tilde{\boldsymbol{f}}(q\boldsymbol{e}_i - q\boldsymbol{e}_j) -\end{aligned}$$

$$\tilde{f}(-q\boldsymbol{e}_i+q\boldsymbol{e}_j)+\tilde{f}(-q\boldsymbol{e}_i-q\boldsymbol{e}_j)]\cdot$$

$$[\tilde{f}(q\boldsymbol{e}_k+q\boldsymbol{e}_l)-\tilde{f}(q\boldsymbol{e}_k-q\boldsymbol{e}_l)-$$

$$\tilde{f}(-q\boldsymbol{e}_k+q\boldsymbol{e}_l)+\tilde{f}(-q\boldsymbol{e}_k-q\boldsymbol{e}_l)]^{\mathrm{T}}+\boldsymbol{Q} \tag{8-40}$$

代入式(8－28)和式(8－36),式(8－40)变换为

$$\begin{aligned}\boldsymbol{P}^{xx}_{n|n-1}=&\frac{1-w_0}{4n_x}\sum_{i=1}^{n_x}\sum_{j=1}^{n_x}[\boldsymbol{f}(\boldsymbol{\chi}^{(i)}_{n-1|n-1})-\boldsymbol{f}(\boldsymbol{\chi}^{(n_x+i)}_{n-1|n-1})]\cdot\\&[\boldsymbol{f}(\boldsymbol{\chi}^{(j)}_{n-1|n-1})-\boldsymbol{f}(\boldsymbol{\chi}^{(n_x+j)}_{n-1|n-1})]^{\mathrm{T}}+\frac{1}{2}\left(\frac{1-w_0}{n_x}\right)^2\cdot\\&\sum_{i=1}^{n_x}\sum_{j=1}^{n_x}[\boldsymbol{f}(\boldsymbol{\chi}^{(i)}_{n-1|n-1})-2\boldsymbol{f}(\boldsymbol{\chi}^{(0)}_{n-1|n-1})+\boldsymbol{f}(\boldsymbol{\chi}^{(n_x+i)}_{n-1|n-1})]\cdot\\&[\boldsymbol{f}(\boldsymbol{\chi}^{(j)}_{n-1|n-1})-2\boldsymbol{f}(\boldsymbol{\chi}^{(0)}_{n-1|n-1})+\boldsymbol{f}(\boldsymbol{\chi}^{(n_x+j)}_{n-1|n-1})]^{\mathrm{T}}+\\&\boldsymbol{Q}+\boldsymbol{HOT}\end{aligned} \tag{8-41}$$

式中:$\boldsymbol{HOT}$ 为高阶交叉项的总和。

8.2.3 多维有限差分观测预测方程

式(7－42)和式(7－43)分别给出了观测值和观测协方差。按照上节使用过的方法,用有限差分逼近取代微分运算,得

$$\hat{\boldsymbol{z}}_{n|n-1}=\sum_{j=0}^{2n_x}w_j\tilde{\boldsymbol{h}}(\boldsymbol{\chi}^{(j)}_{n|n-1}) \tag{8-42}$$

$$\begin{aligned}\boldsymbol{P}^{zz}_{n|n-1}=&\frac{1-w_0}{4n_x}\sum_{j=1}^{n_x}[\boldsymbol{h}(\boldsymbol{\chi}^{(j)}_{n|n-1})-\boldsymbol{h}(\boldsymbol{\chi}^{(n_x+j)}_{n|n-1})]\cdot\\&[\boldsymbol{h}(\boldsymbol{\chi}^{(j)}_{n|n-1})-\boldsymbol{h}(\boldsymbol{\chi}^{(n_x+j)}_{n|n-1})]^{\mathrm{T}}+\\&\frac{1}{2}\left(\frac{1-w_0}{n_x}\right)^2\sum_{j=1}^{n_x}[\boldsymbol{h}(\boldsymbol{\chi}^{(j)}_{n|n-1})-2\boldsymbol{h}(\boldsymbol{\chi}^{(0)}_{n|n-1})+\boldsymbol{h}(\boldsymbol{\chi}^{(n_x+j)}_{n|n-1})]\\&[\boldsymbol{h}(\boldsymbol{\chi}^{(j)}_{n|n-1})-2\boldsymbol{h}(\boldsymbol{\chi}^{(0)}_{n|n-1})+\boldsymbol{h}(\boldsymbol{\chi}^{(n_x+j)}_{n|n-1})]^{\mathrm{T}}+\boldsymbol{R}\end{aligned} \tag{8-43}$$

这里忽略了高阶交叉项,且

$$\boldsymbol{\chi}^{(j)}_{n|n-1}=\hat{\boldsymbol{x}}_{n|n-1}+\sqrt{\frac{n_x}{1-w_0}}\boldsymbol{D}_{n|n-1}\boldsymbol{r}^{(j)} \tag{8-44}$$

由式(7－46),状态观测互协方差预测可写为

$$\boldsymbol{P}^{xz}_{n|n-1}=\boldsymbol{D}_{n|n-1}\left[\sum_{i=1}^{n_x}\frac{\partial}{\partial c_i}\tilde{\boldsymbol{h}}_i(\boldsymbol{c})\right]_{\boldsymbol{c}=\boldsymbol{0}} \tag{8-45}$$

用有限差分等效替换偏导数,式(8 - 45) 可简化为

$$P_{n|n-1}^{xz} = D_{n|n-1}\frac{1}{2q}\sum_{i=1}^{n_x}[\tilde{h}(qe_i) - \tilde{h}(-qe_i)] \tag{8-46}$$

代入式(8 - 36) 和式(8 - 38),式(8 - 46) 变为

$$P_{n|n-1}^{xz} = \frac{1}{2}\sqrt{\frac{1-w_0}{n_x}}D_{n|n-1}\sum_{j=1}^{n_x}[h(\chi_{n|n-1}^{(j)}) - h(\chi_{n|n-1}^{(n_x+j)})] \tag{8-47}$$

8.2.4 多维有限差分卡尔曼滤波器过程

上面推导得的 FDKF 方程包含比线性更高阶的项,因此比 LKF 或线性化 EKF 算法更为准确。如果忽略 HOT 项,FDKF 过程可以概括为表 8 - 2。

像对一维情况的处理那样,这些预测方程也可以通过设 $\omega_0 = 0$ 进行简化,这样可以去掉位于原点的点,得没有自由参数的预测方程。

表 8 - 2 多维有限差分卡尔曼滤波器过程

步骤	内容	公式
步骤(1)	滤波器初始化	设 $0 \leqslant \omega_0 < 1$ 初始化 $\hat{x}_0$ 和 P_0^{xx}
步骤(2)	状态矢量预测	$\chi_{n-1\|n-1}^{(j)} = \hat{x}_{n-1\|n-1} + \sqrt{\frac{n_x}{1-w_0}}D_{n-1\|n-1}r^{(j)}\ j = 0,\cdots,2n_x$ $\hat{x}_{n\|n-1} = \sum_{j=0}^{2n_x} w_j f(\chi_{n-1\|n-1}^{(j)})$ $P_{n\|n-1}^{xx} = \frac{1-w_0}{4n_x}\cdot \sum_{i=1}^{n_x}\sum_{j=1}^{n_x}[f(\chi_{n-1\|n-1}^{(i)}) - f(\chi_{n-1\|n-1}^{(n_x+i)})]\cdot [f(\chi_{n-1\|n-1}^{(j)}) - f(\chi_{n-1\|n-1}^{(n_x+j)})]^{\mathrm{T}} + Q$
步骤(3)	观测值预测	$\chi_{n\|n-1}^{(j)} = \hat{x}_{n\|n-1} + \sqrt{\frac{n_x}{1-w_0}}D_{n\|n-1}r^{(j)}\ j = 0,\cdots,2n_x$ $\hat{z}_{n\|n-1} = \sum_{j=0}^{2n_x} w_j \tilde{h}(\chi_{n\|n-1}^{(j)})$ $P_{n\|n-1}^{zz} = \frac{1-w_0}{4n_x}\sum_{j=1}^{n_x}[h(\chi_{n\|n-1}^{(j)}) - h(\chi_{n\|n-1}^{(n_x+j)})]\cdot [h(\chi_{n\|n-1}^{(j)}) - h(\chi_{n\|n-1}^{(n_x+j)})]^{\mathrm{T}} + R$ $P_{n\|n-1}^{xz} = \frac{1}{2}\sqrt{\frac{1-w_0}{n_x}}D_{n\|n-1}\cdot \sum_{j=1}^{n_x}[h(\chi_{n\|n-1}^{(j)}) - h(\chi_{n\|n-1}^{(n_x+j)})]$

（续）

步骤(4)	卡尔曼滤波器更新	$\boldsymbol{K}_n \triangleq \boldsymbol{P}^{xz}_{n\mid n-1}(\boldsymbol{P}^{zz}_{n\mid n-1})^{-1}$ $\hat{\boldsymbol{x}}_{n\mid n} = \hat{\boldsymbol{x}}_{n\mid n-1} + \boldsymbol{K}_n(z^o_n - \hat{z}_{n\mid n-1})$ $\boldsymbol{P}^{xx}_{n\mid n} = \boldsymbol{P}^{xx}_{n\mid n-1} - \boldsymbol{K}_n \boldsymbol{P}^{zz}_{n\mid n-1}\boldsymbol{K}^{\mathrm{T}}_n$
步骤(5)	存储结果	将 $\hat{\boldsymbol{x}}_{n\mid n}$ 和 $\boldsymbol{P}^{xx}_{n\mid n}$ 存储到跟踪文件中
步骤(6)	返回步骤(2)	

8.3 多维有限差分协方差预测方程的另一种推导

应用有限差分方程(2－73)，由式(7－69)给出的另一种 EKF 状态误差协方差方程变为

$$\begin{aligned}\bar{\boldsymbol{P}}^{xx}_{n\mid n-1} &= \tilde{\boldsymbol{g}}(\boldsymbol{0}) + \frac{1}{2q^2}\sum_{j=1}^{n_x}[\tilde{\boldsymbol{g}}(q\boldsymbol{e}_j) - 2\tilde{\boldsymbol{g}}(\boldsymbol{0}) + \tilde{\boldsymbol{g}}(-q\boldsymbol{e}_j)] + \boldsymbol{Q}\\ &= \left[\frac{q^2 - n_x}{q^2}\right]\tilde{\boldsymbol{g}}(\boldsymbol{0}) + \frac{1}{2q^2}\sum_{j=1}^{n_x}[\boldsymbol{g}(q\boldsymbol{e}_j) + \tilde{\boldsymbol{g}}(-q\boldsymbol{e}_j)] + \boldsymbol{Q}\end{aligned} \quad (8-48)$$

由式(7－67)得 $\tilde{\boldsymbol{g}}(\boldsymbol{c}) \triangleq [\tilde{\boldsymbol{f}}(\boldsymbol{c}) - \hat{\boldsymbol{x}}_{n\mid n-1}][\tilde{\boldsymbol{f}}(\boldsymbol{c}) - \hat{\boldsymbol{x}}_{n\mid n-1}]^{\mathrm{T}}$。利用与上面相同的方法，式(8－48)变换为

$$\bar{\boldsymbol{P}}^{xx}_{n\mid n-1} = \boldsymbol{Q} + \sum_{i=0}^{2n_x} w_j[\boldsymbol{f}(\boldsymbol{\chi}^{(j)}_{n-1\mid n-1}) - \hat{\boldsymbol{x}}_{n\mid n-1}][\boldsymbol{f}(\boldsymbol{\chi}^{(j)}_{n-1\mid n-1}) - \hat{\boldsymbol{x}}_{n\mid n-1}]^{\mathrm{T}} \quad (8-49)$$

和一维情况一样，后一种形式的预测状态协方差矩阵只是二阶偏导。后续章节将表明式(8－37)与式(8－49)对于无迹卡尔曼滤波器来说是相同的。

注意，式(8－43)可以写成另外一种形式，即

$$\bar{\boldsymbol{P}}^{zz}_{n\mid n-1} = \boldsymbol{R} + \sum_{j=0}^{2n_x} w_j[\boldsymbol{h}(\boldsymbol{\chi}^{(j)}_{n\mid n-1}) - \hat{z}_{n\mid n-1}][\boldsymbol{h}(\boldsymbol{\chi}^{(j)}_{n\mid n-1}) - \hat{z}_{n\mid n-1}]^{\mathrm{T}} \quad (8-50)$$

此外，式(8－47)可重写为

$$\bar{\boldsymbol{P}}^{xz}_{n\mid n-1} = \frac{1}{2}\sqrt{\frac{1-w_0}{n_x}}\sum_{j=0}^{2n_x}[\boldsymbol{f}(\boldsymbol{\chi}^{(j)}_{n-1\mid n-1}) - \hat{\boldsymbol{x}}_{n\mid n-1}][\boldsymbol{h}(\boldsymbol{\chi}^{(j)}_{n\mid n-1}) - \hat{z}_{n\mid n-1}]^{\mathrm{T}} \quad (8-51)$$

因为有限差分滤波器与下章将要讨论的无迹卡尔曼滤波器几乎完全相同，本书仅给出 DIFAR 案例分析中无迹卡尔曼滤波器的处理结果。

参考文献

[1] Schei TS. A finite - difference approach to linearization in nonlinear estimation algorithms. *Proc. Am. Control Conf*, 1995,114 - 118.

[2] Schei TS. A finite - difference method for linearization in nonlinear estimation algorithms. *Automatica*, 1997,33(11):2053 - 2058.

[3] N0rgaard M,Poulsen,NK,Ravn O. Advances in Derivative - Free State Estimation for Nonlinear Systems. Technical University of Denmark,IMM - REP - 1998 - 15, 2000.

[4] Norgaard M,Poulsen,NK,Ravn O. New developments state estimation for nonlinear systems. *Automatica*, 2000,36:1627 - 1638.

[5] Ito K,Xiong K. Gaussian filters for nonlinear filtering problems. IEEE *Trans. Automatic Control*, 2000, 45(5):910 - 927.

[6] Wu Y,Hu D,Wu M,Hu X. A numerical - integration perspective on Gaussian filters. *IEEE Trans. Sig. Proc*, 2006,54(8):2910 - 2921.

[7] Lerner UN. Hybrid Bayesian Networks for Reasoning About Complex Systems[D]. Department of Computer Science,Stanford University, October 2002.

[8] McNamee J,Stenger F. Construction of fully symmetric numerical integration formulas. *Numerische Math*, 1967,10:327 - 344.

第9章 Sigma 点类别：无迹卡尔曼滤波器

9.1 单项式容积积分法

在第7章 EKF 推导过程中，所有的非线性函数在某个点处展开为泰勒多项式，并用矩方程来计算展开式中各个单项式的积分。这构成了计算高斯加权非线性函数积分的解析方法。FDKF 继续采用这种解析方法，并用多维 Stirling 插值公式取代泰勒多项式。Stirling 插值逼近要求计算在初始泰勒多项式展开点附近、各维度矢量点均匀分布的非线性函数。尽管 FDKF 和 Sigma 点卡尔曼滤波器形式相同，但它完全用解析的方法得，因此也可以被当作解析线性化卡尔曼滤波器类中的一种。

另一种计算高斯加权积分的方法是利用多维数值积分方法进行计算。在一维情况时，将这些方法归类为著名的（对于了解这类算法的人们来说）正交积分方法。在多维情况时，将它们称为“多维正交”或“容积”积分方法[1,2]。本书中，这两个术语可以互换使用。

第5章式(5-51)~式(5-57)的高斯预测方程为式(5-60)~式(5-62)的一般积分形式。本章将给出一些容积方法来实现数值积分。

对于一般多维高斯加权积分，有

$$\hat{\boldsymbol{g}}(\boldsymbol{x}) = \int \tilde{\boldsymbol{g}}(\boldsymbol{c})\, \mathcal{N}(\boldsymbol{c};\boldsymbol{0},\boldsymbol{I})\,\mathrm{d}\boldsymbol{c} \tag{9-1}$$

由第5章可知，多维高斯加权积分可以用式(5-64)的离散点求和来逼近，即

$$\hat{\boldsymbol{g}}(\boldsymbol{x}) \simeq \sum_{j=1}^{n_s} w_j\, \tilde{\boldsymbol{g}}(\boldsymbol{c}^{(j)}) \tag{9-2}$$

这里 ω_j、$\boldsymbol{c}^{(j)}$ 和 n_s 的具体形式取决于所选的多维容积积分规则。一般来说，n_s 点规则积分计算可表示为

$$R_{n_s} = (w_j, \boldsymbol{c}^{(j)}) \quad j = 1,2,\cdots,1,\cdots,n_s \tag{9-3}$$

这里n_s为积分点的数量，通常与n_x有一定关系。本章以及后面两章里，只考虑单项式规则，下面将给出其定义。

给定矢量$\boldsymbol{c} = [c_1, c_2, \cdots, c_k]^{\mathrm{T}}$和一个多维$d$次单项式（参见第5章结尾部分）

$$a(j_1, j_2, \cdots, j_k) c^{j_1} c^{j_2} \cdots c^{j_k} \tag{9-4}$$

这里$j_1 + j_2 + \cdots + j_k \leqslant d$。例如，继续第5章结尾处的二维例子，对于$\boldsymbol{c} \in \mathbf{R}^2$，由$d(d \leqslant 2)$次单项式扩展成的多项式基为集合$\{1, c_1, c_2, c_1, c_2, c_1^2, c_2^2\}$；对于$d=2$，多项式中的每个单项式必须满足$j_1 + j_2 \leqslant d$条件。一般来说，对于矢量$\boldsymbol{c} \in \mathbf{R}^k$有个$\begin{pmatrix} k+d \\ d \end{pmatrix}$精度为$d$或者更小的不同单项式。二项式系数$\begin{pmatrix} n \\ m \end{pmatrix}$表示不考虑排序情况下，从$n$个不同的物体中取出$m$个可能出现的情况数目。它的定义为

$$\begin{pmatrix} n \\ m \end{pmatrix} \triangleq \frac{n!}{m!\,(n-m)!} \tag{9-5}$$

例9.1 设$k=2$且$d=2$，定义

$$\dim(k,d) \triangleq \begin{pmatrix} k+d \\ d \end{pmatrix} = \begin{pmatrix} 4 \\ 2 \end{pmatrix} = \frac{4!}{2!\,(4-2)!} = 6 \tag{9-6}$$

式(9-3)为n_s点d次单项式法则，由$\sum_{j=1}^{n_s} \omega_j \tilde{\boldsymbol{g}}(\boldsymbol{c}^{(j)})$计算$\int_{\mathbf{R}^k} \tilde{\boldsymbol{g}}(\boldsymbol{c}) \omega(\boldsymbol{c}) \mathrm{d}\boldsymbol{c}$，$\tilde{\boldsymbol{g}}(\boldsymbol{c})$为小于或等于$d$次的多项式。将任意单项式基集合标记为$\{b_1, b_2, \cdots, b_{\dim(k,d)}\}$，用法则$R_{n_s} = \{(\omega_j, \boldsymbol{c}^{(j)}), j = 1, \cdots, n_s\}$对每个次数小于或等于$d$的多项式进行积分，那么必须满足密度加权矩方程的单项式系统[3]

$$\sum_{j=1}^{n_s} w_j b_i(\boldsymbol{c}^{(j)}) = \int_{\mathbf{R}^k} b_i(\boldsymbol{c}) \mathcal{N}(\boldsymbol{c}; \boldsymbol{0}, \boldsymbol{I}) \mathrm{d}\boldsymbol{c} \quad i = 1,2,\cdots,1,\cdots,\dim(k,d) \tag{9-7}$$

对于高斯加权密度，式(9-7)变为

$$\sum_{j=1}^{n_s} w_j b_i(\boldsymbol{c}^{(j)}) = \int_{\mathbf{R}^k} b_i(\boldsymbol{c}) w(\boldsymbol{c}) \mathrm{d}\boldsymbol{c} \quad i = 1,2,\cdots,1,\cdots,\dim(k,d) \tag{9-8}$$

假定想得一个满足高斯密度加权矩（二阶矩）方程的单项式规则，由于$\dim(k,d) = \dim(k,2)$，则有$\{b_1, b_2, \cdots, b_{\dim(k,d)}\} \to \{1, \boldsymbol{c}, \boldsymbol{c}\boldsymbol{c}^{\mathrm{T}}\}$，进而矩方程组则变为

$$\sum_{j=1}^{n_s} w_j = \int_{\mathbf{R}^k} \mathcal{N}(\boldsymbol{c}; \boldsymbol{0}, \boldsymbol{I}) \mathrm{d}\boldsymbol{c} = 1 \tag{9-9}$$

$$\sum_{j=1}^{n_s} w_j \boldsymbol{c}^{(j)} = \int_{\mathbf{R}^k} \boldsymbol{c}\, \mathcal{N}(\boldsymbol{c};\boldsymbol{0},\boldsymbol{I})\,\mathrm{d}\boldsymbol{c} = \boldsymbol{0} \tag{9-10}$$

$$\sum_{j=1}^{n_s} w_j \boldsymbol{c}^{(j)} \boldsymbol{c}^{(j)\mathrm{T}} = \int_{\mathbf{R}^k} \boldsymbol{c}\boldsymbol{c}^{\mathrm{T}}\, \mathcal{N}(\boldsymbol{c};\boldsymbol{0},\boldsymbol{I})\,\mathrm{d}\boldsymbol{c} = \boldsymbol{I} \tag{9-11}$$

积分规则的最后一步是指定一组位于 k 维超球面上的矢量点 $\{\boldsymbol{c}^{(j)}, j=1, 2,\cdots,1,\cdots,n_s\}$，并利用它们通过式(9-9)~式(9-11)求解权值 w_j。

9.2 无迹卡尔曼滤波器

9.2.1 产生背景

无迹卡尔曼滤波器(UKF)最初是由1995年华盛顿州西雅图美国控制会议的两篇论文[4,5]中首次提出。这两篇论文提出了一种新的非线性状态估计线性化逼近方法，与EKF不同，它不需要计算雅可比矩阵。该方法能“有效地通过离散分布(矢量点)逼近高斯分布，允许对每个独立的点应用非线性转换。”特别之处在于，指定一组关于多维高斯分布均值对称的矢量点，将这些选中的点集输入到非线性变换中，利用输出结果点的加权和，可以得更好的变换后均值估计，优于EKF得的估计。

在文献[6-8]中，Julier详尽地阐述了他的算法；在文献[7]中，他将这种方法命名为“分布逼近滤波器”，并给出了与EKF相比，该算法在处理高度非线性函数变换中的优势；在文献[8]中，Julier提出“与任意非线性函数的逼近或转换相比，概率分布函数的逼近更为容易。”然而，在文献[9]中，Lefebvre等给出“对一些回归点，进行统计线性回归，通过线性化过程和测量函数，得出完全相同的估计”。上述方法与本书后面即将给出的完全解析方法有一些相似之处，包括将单项式积分规则用于非线性函数的多项式逼近。下面的方法不涉及任何密度函数逼近，而是对于选定的积分点，计算高斯矩方程组。实践证明，Julier和Uhlmann所提出的Sigma点卡尔曼滤波器是状态矢量仿射变换的必然结果，得的估计方法可用于规范化高斯密度的简单矩积分。

下面介绍无迹卡尔曼滤波器的进一步扩展与推广。参考文献[10]中，Wan和van der Merwe扩展了UKF，使其除状态估计外，还包括参数估计。文献[11]中，提出了一种平方根UKF，可进行状态估计和参数估计。在文献[12]中，将UKF中增加额外的测试点，使得UKF可应用于不连续非线性函数。在文献[13]中，Julier推广了他之前提出的UKF算法，可以将Sigma点缩放至任意维数。在文献[14]中，Tenne和Singh将UKF推广至高阶。针对那些状态(或观测)矢量可以被解析成线性和非线性分量的情况，Briers等在文献[15]中提出一种Rao-Blackwell UKF算法。最近，Banani和Masnadi-Shiraz[16]发表了一种

迭代 UKF 算法。

在 van der Merwe[17]以及 Julier 和 Uhlmann[18]几乎同时发表的论文中，对最初的 UKF 以及其诸多扩展算法做了很好的综述，文中除了给出一些实际应用以及示例之外，还做了非常全面的回顾。

9.2.2 无迹卡尔曼滤波器发展历程

为推导无迹卡尔曼滤波器，假设给定与状态矢量维度相同的矢量点$\{\boldsymbol{c}^{(j)}, j=1,2,\cdots,1,\cdots,n_s\}$，包括构成沿着每根笛卡儿坐标轴，半径为$\pm q$的对称点对，以及位于坐标原点的单个点。除位于坐标原点的点，所有矢量点都分布在同一个n_x维的超球面上。两个矢量点在每一维中的对称性，保证了积分点的总数为$n_s=2n_x+1$。用分布在单位超球面（由式(2-30)得）坐标轴上的单位矢量集$\boldsymbol{r}=[1]\in\mathbf{R}^{n_x}$，以及一个位于坐标原点的点，将矢量点集合表示为$\boldsymbol{c}^{(j)}=q\boldsymbol{r}^{(j)}$。对于$n_x=4$，$\boldsymbol{r}$值由式(8-32)给出。

对于$\boldsymbol{c}\in\mathbf{R}^4$，$n_x=4$，有$n_s=2n_x+1=9$个矢量计算点。利用式(9-9)~式(9-11)给出的矢量点，得如下矩方程组，即

$$\sum_{j=0}^{8} w_j = 1 \tag{9-12}$$

$$\sum_{j=0}^{8} w_j\boldsymbol{c}^{(j)} = q[w_1\boldsymbol{r}^{(1)}+w_2\boldsymbol{r}^{(2)}+\cdots+w_8\boldsymbol{r}^{(8)}]=\boldsymbol{0} \tag{9-13}$$

$$\sum_{j=0}^{8} w_j\boldsymbol{c}^{(j)}\boldsymbol{c}^{(j)\mathrm{T}} = q^2[w_1\boldsymbol{r}^{(1)}\boldsymbol{r}^{(1)\mathrm{T}}+\cdots+w_8\boldsymbol{r}^{(8)}\boldsymbol{r}^{(8)\mathrm{T}}]=\boldsymbol{I} \tag{9-14}$$

式(9-13)可重写为

$$(w_1-w_5)\begin{bmatrix}1\\0\\0\\0\end{bmatrix}+(w_2-w_6)\begin{bmatrix}0\\1\\0\\0\end{bmatrix}+(w_3-w_7)\begin{bmatrix}0\\0\\1\\0\end{bmatrix}+(w_4-w_8)\begin{bmatrix}0\\0\\0\\1\end{bmatrix}=\begin{bmatrix}0\\0\\0\\0\end{bmatrix}$$

这可以通过设置$\omega_1=\omega_5$，$\omega_2=\omega_6$，$\omega_3=\omega_7$和$\omega_4=\omega_8$来实现。

式(9-14)可写为

$$(w_1+w_5)q^2\begin{bmatrix}1&0&0&0\\0&0&0&0\\0&0&0&0\\0&0&0&0\end{bmatrix}+(w_2+w_6)q^2\begin{bmatrix}0&0&0&0\\0&1&0&0\\0&0&0&0\\0&0&0&0\end{bmatrix}+$$

$$(w_3+w_7)q^2\begin{bmatrix}0&0&0&0\\0&0&0&0\\0&0&1&0\\0&0&0&0\end{bmatrix}+(w_4+w_6)q^2\begin{bmatrix}0&0&0&0\\0&0&0&0\\0&0&0&0\\0&0&0&1\end{bmatrix}$$

$$=\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{bmatrix}\tag{9-15}$$

进而得方程组

$$(w_1+w_5)q^2=2w_1q^2=1\longrightarrow q=\sqrt{\frac{1}{2w_1}}\tag{9-16}$$

$$(w_2+w_6)q^2=2w_2q^2=1\longrightarrow q=\sqrt{\frac{1}{2w_2}}\tag{9-17}$$

因此有 $\omega_1=\omega_2$。同理得

$$w_j=w_1\quad j=2,3,\cdots,8\tag{9-18}$$

由式(9－12)可知

$$w_0+8w_1=1\tag{9-19}$$

则有

$$w_1=\frac{1-w_0}{8}\tag{9-20}$$

将式(9－20)代入式(9－16)，得

$$q=\sqrt{\frac{4}{1-w_0}}\tag{9-21}$$

可以推导出

$$\boldsymbol{c}^{(j)}=q\boldsymbol{r}^{(j)}=\sqrt{\frac{4}{1-w_0}}\boldsymbol{r}^{(j)}\quad j=0,1,\cdots,8\tag{9-22}$$

$$w_j=\begin{cases}w_0 & j=0\\ \dfrac{1-w_0}{8} & j=1,2,\cdots,8\end{cases}\tag{9-23}$$

式中：ω_0 为自由参数，为满足式(9－12)且保证所有的权重均为正值，需限制$0\leqslant\omega_0\leqslant1$。

对于 $\boldsymbol{c}\in\mathbf{R}^{n_x}$的一般情况，通过数学归纳法，有

$$\boldsymbol{c}^{(j)}=q\boldsymbol{r}^{(j)}=\sqrt{\frac{n_x}{1-w_0}}\boldsymbol{r}^{(j)}\quad j=0,1,\cdots,2n_x\tag{9-24}$$

$$w_j=\begin{cases}w_0 & j=0\\ \dfrac{1-w_0}{2n_x} & j=1,\cdots,2n_x\end{cases}\tag{9-25}$$

9.2.3 无迹卡尔曼滤波器状态矢量预测方程

由式(5－51)、式(9－1)和式(9－2),可将状态矢量预测方程写为

$$\hat{\boldsymbol{x}}_{n|n-1}=\int\tilde{\boldsymbol{f}}(\boldsymbol{c})\,\mathcal{N}(\boldsymbol{c};\boldsymbol{0},\boldsymbol{I})\mathrm{d}\boldsymbol{c}=\sum_{j=0}^{2n_x}w_j\tilde{\boldsymbol{f}}(\boldsymbol{c}^{(j)})\tag{9-26}$$

利用式(5－53),矢量积分点 $\boldsymbol{c}^{(j)}$ 可写为

$$\tilde{\boldsymbol{f}}(\boldsymbol{c}^{(j)})=\boldsymbol{f}(\hat{\boldsymbol{x}}_{n-1|n-1}+\boldsymbol{D}_{n-1|n-1}\boldsymbol{c}^{(j)})\quad j=0,1,\cdots,2n_x\tag{9-27}$$

将 UKF 中 Sigma 点定义为

$$\boldsymbol{\chi}_{n-1|n-1}^{(j)}\triangleq\hat{\boldsymbol{x}}_{n-1|n-1}+\boldsymbol{D}_{n-1|n-1}\boldsymbol{c}^{(j)}\quad j=0,1,\cdots,2n_x\tag{9-28}$$

式(9－26)状态矢量预测方程则变为

$$\hat{\boldsymbol{x}}_{n|n-1}=\sum_{j=0}^{2n_x}w_j\boldsymbol{f}(\boldsymbol{\chi}_{n-1|n-1}^{(j)})\tag{9-29}$$

这里 $\boldsymbol{c}^{(j)}$ 和 ω_j 分别由式(9－24)和式(9－25)确定,且有 $\boldsymbol{D}_{n-1|n-1}=[\boldsymbol{P}_{n-1|n-1}^{xx}]^{1/2}$。

9.2.4 无迹卡尔曼滤波器状态矢量协方差预测方程

令

$$\tilde{\boldsymbol{g}}(\boldsymbol{c})\triangleq[\tilde{\boldsymbol{f}}(\boldsymbol{c})-\hat{\boldsymbol{x}}_{n|n-1}][\tilde{\boldsymbol{f}}(\boldsymbol{c})-\hat{\boldsymbol{x}}_{n|n-1}]^{\mathrm{T}}\tag{9-30}$$

将其代入式(9－26)替换 $\tilde{\boldsymbol{f}}(\boldsymbol{c})$,则式(9－29)变为如下的状态协方差预测方程,即

$$\begin{aligned}\boldsymbol{P}_{n|n-1}^{xx}&=\int\tilde{\boldsymbol{g}}(\boldsymbol{c})\,\mathcal{N}(\boldsymbol{c};\boldsymbol{0},\boldsymbol{I})\mathrm{d}\boldsymbol{c}+\boldsymbol{Q}=\sum_{j=0}^{2n_x}w_j\tilde{\boldsymbol{g}}(\boldsymbol{c}^{(j)})+\boldsymbol{Q}\\&=\sum_{j=0}^{2n_x}w_j[\boldsymbol{f}(\boldsymbol{\chi}_{n-1|n-1}^{(j)})-\hat{\boldsymbol{x}}_{n|n-1}]\cdot[\boldsymbol{f}(\boldsymbol{\chi}_{n-1|n-1}^{(j)})-\hat{\boldsymbol{x}}_{n|n-1}]^{\mathrm{T}}+\boldsymbol{Q}\end{aligned}\tag{9-31}$$

9.2.5 无迹卡尔曼滤波器观测值预测方程

利用相同的方法,观测值预测方程推导为

$$\hat{\boldsymbol{z}}_{n|n-1}=\sum_{j=0}^{2n_x}w_j\boldsymbol{h}(\boldsymbol{\chi}_{n|n-1}^{(j)})\tag{9-32}$$

$$\boldsymbol{P}_{n|n-1}^{zz}=\sum_{j=0}^{2n_x}w_j[\boldsymbol{h}(\boldsymbol{\chi}_{n|n-1}^{(j)})-\hat{\boldsymbol{z}}_{n|n-1}][\boldsymbol{h}(\boldsymbol{\chi}_{n|n-1}^{(j)})-\hat{\boldsymbol{z}}_{n|n-1}]^{\mathrm{T}}+\boldsymbol{R}\tag{9-33}$$

$$P_{n|n-1}^{xz} = \sum_{j=0}^{2n_x} w_j [f(\chi_{n-1|n-1}^{(j)}) - \hat{x}_{n|n-1}] \cdot [h(\chi_{n|n-1}^{(j)}) - \hat{z}_{n|n-1}]^T \quad (9-34)$$

$$\chi_{n|n-1}^{(j)} \triangleq \hat{x}_{n|n-1} + D_{n|n-1} c^{(j)}, j = 0,1,\cdots,2n_x \quad (9-35)$$

注意:UKF 状态、观测矢量预测方程式(9-29)和式(9-32)分别与有限差分预测方程式(8-37)和式(8-42)的形式相同,而 UKF 协方差预测方程式(9-31)、式(9-33)和式(9-34)分别与有限差分方程的第二种形式(8-51)~式(8-49)相同。

9.2.6 无迹卡尔曼滤波器过程

上面推导得的 UKF 方程包含比线性更高阶的项,因此比 LKF 或线性化 EKF 算法更为准确。表 9-1 给出了 UKF 过程的总结。

表 9-1 多维无迹卡尔曼滤波过程

步骤	内容	公式
步骤(1)	滤波器初始化	设 $0 \leqslant w_0 < 1$ 初始化和 $\hat{x}_0$ 和 P_0^{xx}
步骤(2)	状态矢量预测	$\chi_{n-1|n-1}^{(j)} = \hat{x}_{n-1|n-1} + D_{n-1|n-1} c^{(j)} \quad j = 0,\cdots,2n_x$ $\hat{x}_{n|n-1} = \sum_{j=0}^{2n_x} w_j f(\chi_{n-1|n-1}^{(j)})$ $P_{n|n-1}^{xx} = \sum_{j=0}^{2n_x} w_j [f(\chi_{n-1|n-1}^{(j)}) - \hat{x}_{n|n-1}] \cdot [f(\chi_{n-1|n-1}^{(j)}) - \hat{x}_{n|n-1}]^T +$ Q
步骤(3)	观测值预测	$\chi_{n|n-1}^{(j)} = \hat{x}_{n|n-1} + D_{n|n-1} c^{(j)} \quad j = 0,\cdots,2n_x$ $\hat{z}_{n|n-1} = \sum_{j=0}^{2n_x} w_j \tilde{h}(\chi_{n|n-1}^{(j)})$ $P_{n|n-1}^{zz} = \sum_{j=0}^{2n_x} w_j [h(\chi_{n|n-1}^{(j)}) - \hat{z}_{n|n-1}]$ $[h(\chi_{n|n-1}^{(j)}) - \hat{z}_{n|n-1}]^T + R$ $P_{n|n-1}^{xz} = \sum_{j=0}^{2n_x} w_j [f(\chi_{n-1|n-1}^{(j)}) - \hat{x}_{n|n-1}] \cdot [h(\chi_{n|n-1}^{(j)}) - \hat{z}_{n|n}]^T$
步骤(4)	卡尔曼滤波器更新	$K_n \triangleq P_{n|n-1}^{xz} (P_{n|n-1-1}^{zz})^{-1}$ $\hat{x}_{n|n} = \hat{x}_{n|n-1} + K_n (z_n^o - \hat{z}_{n|n-1})$ $P_{n|n}^{xx} = P_{n|n-1}^{xx} - K_n P_{n|n-1}^{zz} K_n^T$
步骤(5)	存储结果	将 $\hat{x}_{n|n}$ 和 $P_{n|n}^{xx}$ 存储到跟踪文件中
步骤(6)	返回步骤(2)	

9.2.7 无迹卡尔曼滤波器的另一个版本

从式(9-12)之前内容可知,除原点处的点之外,所有的点都分布在一个半径为 q 的超球面上。由于原点处的点可以任意添加,取 $\omega_0 = 0$,从积分和中移除位于原点的 Sigma 点,不会损失其一般性。注意,位于原点处的 Sigma 点(位于

$r^{(0)}$处）与$\hat{x}_{n-1|n-1}$或$\hat{x}_{n|n-1}$相同。这简化了 UKF，消除了选取一个满足$0 \leq \omega_0 \leq 1$条件的值的任意性。这样所有的求和运算具有如下限制：对于$j=1,\cdots,2n_x$，有$\omega_j = 1/2n_x$；$\forall j$，且$\chi^{(j)} = \hat{x} + \sqrt{n_x} D r^{(j)}$。

在文献[2,20]首次提出球面径向容积准则方法之后，Arasaratnam 和 Haykin[19]给出了相同的结果，他们采用球面径向容积准则，将笛卡儿坐标下的高斯加权积分变换到极坐标下。在推导 UKF 积分准则的过程中，式(9-12)~式(9-14)需要求解$2n_x+1$个联立方程。在球面径向法中，利用高斯加权积分的完全对称性质，将联立方程的数量减少为 2 个。但这个方法与上面推导的 UKF 在$\omega_0=0$时的结果相同，都是非线性卡尔曼滤波器。

9.3 UKF 在 DIFAR 船舶跟踪的应用案例分析

由于在 DIFAR 跟踪问题中，动态方程为线性的，表 9-2 给出了解决这个问题所采用的 UKF 过程。注意，这个过程对于所有的 Sigma 点卡尔曼滤波器通用，唯一的区别是$\{(\omega_j, c^{(j)}), j=1,\cdots,n_s\}$的取值。对于 DIFAR 实例，在表 9-2 中，$n_x=4$，取$r=[1]\in \mathbf{R}^4$，加上位于原点的 Sigma 点，UKF 中共有 9 个 Sigma 点。对于 Sigma 点，$\{(\omega_j, c^{(j)}), j=1,\cdots,n_s\}$的取值由式(9-22)和式(9-23)决定，这里取$\omega_0=0.25$。

表 9-2　应用于 DIFAR 跟踪的多维 Sigma 点卡尔曼滤波器过程

步骤(1)	滤波器初始化	设$0 \leq \omega_0 < 1$ 初始化和$\hat{x}_0$和P_0^{xx}
步骤(2)	状态矢量预测	$\hat{x}_{n\|n-1} = F\hat{x}_{n-1\|n-1}$ $P_{n\|n-1}^{xx} = FP_{n-1\|n-1}^{xx}F^{\mathrm{T}} + Q$ $D_{n\|n-1} = [P_{n\|n-1}^{xx}]^{1/2}$
步骤(3)	观测值预测	$\chi_{n\|n-1}^{(j)} = \hat{x}_{n\|n-1} + D_{n\|n-1}c^{(j)}, j=0,\cdots,2n_x$ $\hat{z}_{n\|n-1} = \sum_{j=0}^{2n_x} w_j h(\chi_{n\|n-1}^{(j)})$ $P_{n\|n-1}^{zz} = \sum_{j=0}^{2n_x} w_j [h(x_{n\|n-1}^{(j)}) - \hat{z}_{n\|n-1}][h(\chi_{n\|n-1}^{(j)}) - \hat{z}_{n\|n-1}]^{\mathrm{T}} + R$ $P_{n\|n-1}^{xz} = \sum_{j=0}^{2n_x} w_j [f(\chi_{n-1\|n-1}^{(j)}) - \hat{x}_{n\|n-1}]\cdot[h(\chi_{n\|n-1}^{(j)}) - \hat{z}_{n\|n-1}]^{\mathrm{T}}$
步骤(4)	卡尔曼滤波器更新	$K_n \triangleq P_{n\|n-1}^{xz}(P_{n\|n-1}^{zz})^{-1}$ $\hat{x}_{n\|n} = \hat{x}_{n\|n-1} + K_n(z_n^o - \hat{z}_{n\|n-1})$ $P_{n\|n}^{xx} = P_{n\|n-1}^{xx} - K_n P_{n\|n-1}^{zz} K_n^{\mathrm{T}}$
步骤(5)	存储结果	将$\hat{x}_{n\|n}$和$P_{n\|n}^{xx}$存储到跟踪文件中
步骤(6)	返回步骤(2)	

在这个 UKF 滤波器中,采用了与 EKF 相同的初始化以及动态噪声值 q,具体参见 7.4.3 节和 7.4.4 节内容。

在 DIFAR 实例中,输入 6 种不同信噪比的信号,采用 UKF 跟踪器进行跟踪,与图 7-3 给出的 EKF 跟踪图做对比,比较 100 次蒙特卡洛输出轨迹。结果表明,在不同信噪比条件下,两种跟踪算法的输出结果在发散航迹的数目上几乎相同,所以这里不再给出 UKF 的跟踪结果比较图。

参考文献

[1] Lu J, Darmofal D L. Higher - dimensional integration with Gaussian weights for applications in probabilistic design. *SIAM J. Sci. Comput*,2004,26(2):613-624.

[2] Dunavant D A. Efficient symmetrical cubature rules for complete polynomials of higher degree over the unit cube. *Int. J. Numer. Methods Eng*,1986,23:397-407.

[3] Haber S. Numerical evaluation of multiple integrals. *SIAM Rev*,1970,12(4):481-526.

[4] Quine B, Uhlmann J, Durrant - Whyte H. Implicit Jacobians for linearized state estimation in nonlinear systems. *Proc. Am. Control Conf*,1995:1645-1646.

[5] Julier SJ, Uhlmann JK, Durrant - Whyte HF. A new approach for filtering nonlinear systems. Proc. Am. Control Conf, 1995:1628-1632.

[6] Julier S, Uhlmann, JK. A General Method for Approximating Nonlinear Transformations of Probability Distributions. Robotics Research Group, Department of Engineering Science, University of Oxford, UK, 1996.

[7] Julier SJ. Process Models for the Navigation of High Speed Land Vehicles, Dissertation. Oxford. UK: Wadham College,1997.

[8] Julier S, Uhlmann J, Durrant - Whyte HF. A new method for the nonlinear transformation of means and covariances in filters and estimators. *IEEE Trans. Automatic Control*,2000,45(3):477-482.

[9] Lefebvre T, Bruyninckx H, De Schutter J. Comments on "A new method for the nonlinear transformation of means and covariances in filters and estimators." *IEEE Trans. Automatic Control*, 2002, 47 (8): 1406-1408.

[10] Wan EA, van der Merwe R. The Unscented Kalman Filter for Nonlinear Estimation. *Adaptive Systems for Signal Processing, Communications, and Control Symposium, AS - SPCC*,2000:153-158.

[11] van der Merwe R, Wan EA. The Square - Root Unscented Kalman Filter for State and Parameter Estimation. ICASSP 01, Vol. 6, 2001:3461-3464.

[12] Van Zandt J. A more robust unscented transform. *Proc. SPIE*,2001,4473:371-379.

[13] Julier SJ. The scaled unscented transform. *Proc. Am. Control Conf*,2002:4555-4559.

[14] Tenne D, Singh T. The higher order unscented filter. *Proc. Am. Control Conf*,2003,3:2441-2446.

[15] Briers M, Maskell SR. Wright R. A Rao - Blackwellised unscented Kalman filter. *Inform. Fusion*,2003, 1:55-61.

[16] Banani SA, Masnadi - Shirazi MA. A new version of unscented Kalman filter. *Proc. World Acad. Sci. Eng. Technol*,2007,20:192-197.

[17] van der Merwe R. Sigma - Point Kalman Filters for Probabilistic Inference in Dynamic State - Space Mod-

els, Dissertation. Portland OR: Oregon Health & Science University, 2004.

[18] Julier SJ, Uhlmann JK. Unscented filtering and nonlinear estimation. *Proc. IEEE*, 2004, 92 (3): 401 - 422.

[19] Arasaratnam I, Haykin S. Cubature Kalman filters. *IEEE Trans. Automatic Control*, 2009, 54(6): 1254 - 1269.

[20] Monahan J, Genx A. Spherical - radial integration rules for Bayesian computation. *J. Am. Stat. Assoc*, 1997, 92(438): 664 - 674.

第 10 章 Sigma 点类别:超球面单形卡尔曼滤波器

对于 UKF 算法,除了在原点处的采样点,矩积分方程(9-9)~式(9-11)用到的矢量采样点关于 **0** 对称,与原点距离相等,落在 n_x 维笛卡儿坐标系的坐标轴上。对于超球面单形卡尔曼滤波器,要求关于原点对称替换为要求所有矢量点与原点及彼此间的距离相等,因此分布在一个 n_x 维单形(图 2-4)的顶点上。这一采样规则由 Julier[1]首次提出。

将采样(Sigma)点重新定义为 $\boldsymbol{c}^{(j)} \to \boldsymbol{c}_j$,所有单形矢量点 $\boldsymbol{c}_j$ 都必须满足矩方程式(9-9)~式(9-11)。为清晰起见,这里再重复一下,即

$$\sum_{j=0}^{n_x} w_j = 1 \tag{10-1}$$

$$\sum_{j=0}^{n_x} w_j \boldsymbol{c}_j = \boldsymbol{0} \tag{10-2}$$

$$\sum_{j=0}^{n_x} w_j \boldsymbol{c}_j \boldsymbol{c}_j^{\mathrm{T}} = \boldsymbol{I} \tag{10-3}$$

10.1 一维超球面单形 Sigma 点

考虑一维情况($n_x = 1$),在一条直线上取三个任意点$(c_0, c_1, c_2) = (0, q_1, -q_2)$,如图 2-4 所示。将这些点代入矩方程中,得下面的方程组,即

$$w_0 + w_1 + w_2 = 1 \tag{10-4}$$

$$w_1 q_1 - w_2 q_2 = 0 \tag{10-5}$$

$$w_1 q_1^2 + w_2 q_2^2 = 1 \tag{10-6}$$

注意:矢量点必须与坐标原点等距。令 $q_2 = q_1$,由式(10-5),有

$$w_2 = w_1 \tag{10-7}$$

由式(10-4),得

$$w_1 = w_2 = \frac{1 - w_0}{2} \tag{10-8}$$

根据式(10-6),有

$$q_1 = \frac{1}{\sqrt{2w_1}} \tag{10-9}$$

因此,对于一维情况,Sigma 点可定义为

$$c_j^{1D} = \begin{cases} 0 & j = 0 \\ 1/\sqrt{2w_1} & j = 1 \\ -1/\sqrt{2w_1} & j = 2 \end{cases} \tag{10-10}$$

权值分别为

$$0 \leqslant w_0 \leqslant 1 \tag{10-11}$$

$$w_1 = \frac{1 - w_0}{2} \tag{10-12}$$

$$w_2 = \frac{1 - w_0}{2} \tag{10-13}$$

一维情况下有 $x = \hat{x} + \sigma c$,得一维 Sigma 点为

$$\chi_0 = \hat{x} \tag{10-14}$$

$$\chi_1 = \hat{x} + \sigma c_1^{1D} \tag{10-15}$$

$$\chi_2 = \hat{x} + \sigma c_2^{1D} \tag{10-16}$$

10.2 二维超球面单形 Sigma 点

图 10-1 给出了二维情况下位于单形顶点的矢量点示例。在这个等边三角形中,每条边与角都是相等的,所有的采样点与位于原点的点距离相等。在二维情况下,矢量采样点集合为$(\boldsymbol{c}_0, \boldsymbol{c}_1, \boldsymbol{c}_2, \boldsymbol{c}_3) = \{(0,0)^{\mathrm{T}}, (q_1 - q_3)^{\mathrm{T}}, (-q_1, -q_3)^{\mathrm{T}}, (0, s_3 q_3)^{\mathrm{T}}\}$。这里,$s_3$ 是一个尺度参数,其值将在后面求解。

再次强调,必须满足矩方程(10-1)矩方程(10-3)。由式(10-1),有

$$w_0 + w_1 + w_2 + w_3 = 1 \tag{10-17}$$

由式(10-2),得

$$w_1 q_1 - w_2 q_1 = 0 \tag{10-18}$$

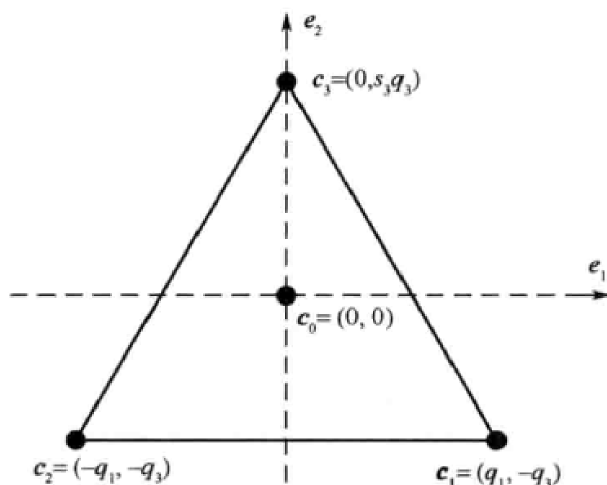


图 10－1　由 4 个矢量采样点构成的二维单形

$$-w_1q_3-w_2q_3+w_3s_3q_3=0 \tag{10－19}$$

由式(10－3),得

$$w_1\begin{bmatrix}q_1^2 & 0\\0 & q_3^2\end{bmatrix}+w_2\begin{bmatrix}q_1^2 & 0\\0 & q_3^2\end{bmatrix}+w_3\begin{bmatrix}0 & 0\\0 & s_3^2q_3^2\end{bmatrix}=\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix} \tag{10－20}$$

可简化方程组为

$$w_1q_1^2+w_2q_1^2=1 \tag{10－21}$$

$$w_1q_3^2+w_2q_3^2+w_3s_3^2q_3^2=1 \tag{10－22}$$

由式(10－18),有

$$w_2=w_1 \tag{10－23}$$

且式(10－22)变为

$$2w_1q_3^2+w_3s_3^2q_3^2=1 \tag{10－24}$$

因为每一个点与原点的距离相等,可以假定每个点的权重是相同的 $w_3=w_1$。由式(10－17),有

$$w_1=w_2=w_3=\frac{1-w_0}{3} \tag{10－25}$$

由式(10－19)可知 $s_3=2$,这样式(10－24)可写为

$$q_3=\frac{1}{\sqrt{6w_1}}=\frac{1}{\sqrt{2(2+1)w_1}} \tag{10－26}$$

综上,在二维情况下,有如下采样点

$$\boldsymbol{c}_0^{2D}=\begin{bmatrix}0\\0\end{bmatrix}=\begin{bmatrix}c_0^{1D}\\0\end{bmatrix} \tag{10－27}$$

$$\boldsymbol{c}_1^{2D} = \begin{bmatrix} c_1^{1D} \\ \dfrac{1}{\sqrt{2(2+1)w_1}} \end{bmatrix} \tag{10-28}$$

$$\boldsymbol{c}_2^{2D} = \begin{bmatrix} c_2^{1D} \\ -\dfrac{1}{\sqrt{2(2+1)w_1}} \end{bmatrix} \tag{10-29}$$

$$\boldsymbol{c}_3^{2D} = \begin{bmatrix} c_0^{1D} \\ \dfrac{2}{\sqrt{2(2+1)w_1}} \end{bmatrix} \tag{10-30}$$

权值为

$$0 \leqslant w_0 \leqslant 1 \tag{10-31}$$

$$w_j = \frac{1-w_0}{3} \quad j = 1,2,3 \tag{10-32}$$

则式(10－28)和式(10－29)可合并为一个方程

$$\boldsymbol{c}_j^{2D} = \begin{bmatrix} c_j^{1D} \\ -\dfrac{1}{\sqrt{2(2+1)w_1}} \end{bmatrix} \quad j = 1,2 \tag{10-33}$$

10.3 高维超球面单形 Sigma 点

对于一般的 n_x 维情况，则有

$$\boldsymbol{c}_0^{n_x D} = \boldsymbol{0}_{n_x} = \begin{bmatrix} \boldsymbol{c}_0^{(nx-1)D} \\ 0 \end{bmatrix}, \quad j = 0 \tag{10-34}$$

$$\boldsymbol{c}_j^{n_x D} = \begin{bmatrix} \boldsymbol{c}_j^{(n_x-1)D} \\ -\dfrac{1}{\sqrt{n_x(n_x+1)w_1}} \end{bmatrix}, \quad j = 1,\cdots,n_x \tag{10-35}$$

$$\boldsymbol{c}_{n_x+1}^{n_x D} = \begin{bmatrix} \boldsymbol{c}_0^{(n_x-1)D} \\ \dfrac{n_x}{\sqrt{n_x(n_x+1)w_1}} \end{bmatrix} \tag{10-36}$$

权值为

$$0 \leqslant w_0 \leqslant 1$$

$$w_j = \frac{1 - w_0}{n_x + 1} \quad j = 1, \cdots, n_x + 1 \tag{10-37}$$

由于 $\boldsymbol{x} = \hat{\boldsymbol{x}} + \boldsymbol{Dc}$，采样点成为超球面单形 Sigma 点，即

$$\boldsymbol{\chi}^{(j)} = \hat{\boldsymbol{x}} + \boldsymbol{D}\boldsymbol{c}_j^{n_x\mathrm{D}} \quad j = 0, \cdots, n_x + 1 \tag{10-38}$$

10.4 超球面单形卡尔曼滤波器

式(5-51)现在可以写为

$$\hat{\boldsymbol{x}}_{n|n-1} = \int \tilde{\boldsymbol{f}}(\boldsymbol{c})\, \mathcal{N}(\boldsymbol{c}; \boldsymbol{0}, \boldsymbol{I})\,\mathrm{d}\boldsymbol{c} = \sum_{j=0}^{n_x+1} w_j \boldsymbol{f}(\boldsymbol{\chi}_{n-1|n-1}^{(j)}) \tag{10-39}$$

由式(10-38)，可以确定如下 Sigma 点，即

$$\boldsymbol{\chi}_{n-1|n-1}^{(j)} = \hat{\boldsymbol{x}}_{n-1|n-1} + \boldsymbol{D}_{n-1|n-1}\boldsymbol{c}_j^{n_x\mathrm{D}} \quad j = 0, \cdots, n_x + 1 \tag{10-40}$$

其中 $\boldsymbol{c}_j^{n_x\mathrm{D}}$ 由式(10-34)~式(10-36)定义。运用与 UKF 类似方法，可以得其余的预测方程。其中，状态协方差为

$$\begin{aligned}\boldsymbol{P}_{n|n-1}^{xx} &= \int \tilde{\boldsymbol{g}}(\boldsymbol{c})\, \mathcal{N}(\boldsymbol{c}; \boldsymbol{0}, \boldsymbol{I})\,\mathrm{d}\boldsymbol{c} = \sum_{j=0}^{n_x+1} w_j \tilde{\boldsymbol{g}}(\boldsymbol{c}^{(j)}) \\ &= \sum_{j=0}^{n_x+1} w_j [\boldsymbol{f}(\boldsymbol{\chi}_{n-1|n-1}^{(j)}) - \hat{\boldsymbol{x}}_{n|n-1}] \cdot [\boldsymbol{f}(\boldsymbol{\chi}_{n-1|n-1}^{(j)}) - \hat{\boldsymbol{x}}_{n|n-1}]^{\mathrm{T}} + \boldsymbol{Q}\end{aligned} \tag{10-41}$$

观测值预测方程为

$$\hat{\boldsymbol{z}}_{n|n-1} = \sum_{j=0}^{n_x+1} w_j \boldsymbol{h}(\boldsymbol{\chi}_{n|n-1}^{(j)}) \tag{10-42}$$

$$\boldsymbol{P}_{n|n-1}^{zz} = \sum_{j=0}^{n_x+1} w_j [\boldsymbol{h}(\boldsymbol{\chi}_{n|n-1}^{(j)}) - \hat{\boldsymbol{z}}_{n|n-1}] \cdot \tag{10-43}$$

$$[\boldsymbol{h}(\boldsymbol{\chi}_{n|n-1}^{(j)}) - \hat{\boldsymbol{z}}_{n|n-1}]^{\mathrm{T}} + \boldsymbol{R} \tag{10-44}$$

$$\boldsymbol{P}_{n|n-1}^{xz} = \sum_{j=0}^{n_x+1} w_j [\boldsymbol{f}(\boldsymbol{\chi}_{n-1|n-1}^{(j)}) - \hat{\boldsymbol{x}}_{n|n-1}] \cdot [\boldsymbol{h}(\boldsymbol{\chi}_{n|n-1}^{(j)}) - \hat{\boldsymbol{z}}_{n|n-1}]^{\mathrm{T}} \tag{10-45}$$

$$\boldsymbol{\chi}_{n|n-1}^{(j)} = \hat{\boldsymbol{x}}_{n|n-1} + \boldsymbol{D}_{n|n-1}\boldsymbol{c}_j^{n_x\mathrm{D}} \quad j = 0, \cdots, n_x + 1 \tag{10-46}$$

10.5 超球面单形卡尔曼滤波器过程

表 10-1 概括了 SSKF 过程，注意表 10-1、表 8-2 和表 9-1 之间的唯一区别是求和的上限。由于 Sigma 点卡尔曼滤波器之间的相似性，后续章节把所有的Sigma点卡尔曼滤波器合并成一个通用的过程。

表 10－1　多维超球面单形卡尔曼滤波器过程

步骤(1)	滤波器初始化	设 $0 \leqslant \omega_0 < 1$ 初始化和 $\hat{x}_0$ 和 P_0^{xx}
步骤(2)	状态矢量预测	$\chi_{n-1\|n-1}^{(j)} = D_{n-1\|n-1} c_j^{n_x D} \quad j = 0, \cdots, n_x + 1$ $\hat{x}_{n\|n-1} = \sum_{j=0}^{n_x+1} w_j f(\chi_{n-1\|n-1}^{(j)})$ $P_{n\|n-1}^{xx} = \sum_{j=0}^{n_x+1} w_j [f(\chi_{n-1\|n-1}^{(j)}) - \hat{x}_{n\|n-1}] \cdot [f(\chi_{n-1\|n-1}^{(j)}) - \hat{x}_{n\|n-1}]^{\mathrm{T}} + Q$
步骤(3)	观测值预测	$\chi_{n\|n-1}^{(j)} = \hat{x}_{n\|n-1} + D_{n\|n-1} c_j^{n_x D} \quad j = 0, \cdots, 2n_x$ $\hat{z}_{n\|n-1} = \sum_{j=0}^{n_x+1} w_j h(\chi_{n\|n-1}^{(j)})$ $P_{n\|n-1}^{zz} = \sum_{j=0}^{n_x+1} w_j [h(\chi_{n\|n-1}^{(j)}) - \hat{z}_{n\|n-1}][h(\chi_{n\|n-1}^{(j)}) - \hat{z}_{n\|n-1}]^{\mathrm{T}} + R$ $P_{n\|n-1}^{xz} = \sum_{j=0}^{n_x+1} w_j [f(\chi_{n-1\|n-1}^{(j)}) - \hat{x}_{n\|n-1}] \cdot [h(\chi_{n\|n-1}^{(j)}) - \hat{z}_{n\|n-1}]^{\mathrm{T}}$
步骤(4)	卡尔曼滤波器更新	$K_n \triangleq P_{n\|n-1}^{xz} (P_{n\|n-1}^{zz})^{-1}$ $\hat{x}_{n\|n} = \hat{x}_{n\|n-1} + K_n (z_n^o - \hat{z}_{n\|n-1})$ $P_{n\|n}^{xx} = P_{n\|n-1}^{xx} - K_n P_{n\|n-1}^{zz} K_n^{\mathrm{T}}$
步骤(5)	存储结果	将 $\hat{x}_{n\|n}$ 和 $P_{n\|n}^{xx}$ 存储到跟踪文件中
步骤(6)	返回步骤(2)	

10.6　SSKF 在 DIFAR 船舶跟踪的应用案例分析

因为 DIFAR 跟踪问题的动态方程是线性的，求解这个问题的 SSKF 过程与表 9－2 中的 UKF 过程除了求和的上限，其余均相同。

对于 DIFAR 情况，取 $n_x = 4$，$c^{(j)} \to c_j^{4D}$。因为包含了位于原点的 Sigma 点，所以 SSKF 有 6 个 Sigma 点。对于 Sigma 点，$\{(\omega_j, c_j^{4D}), j = 0, 1, \cdots, 5\}$ 值由式(10－34)～式(10－37)给出，这里取 $\omega_0 = 0.25$。在这个 SSKF 滤波器中，采用了与 EKF 相同的初始化以及动态噪声值 q(分别在 7.4.3 节和 7.4.4 节中给出)。比较 100 次蒙特卡洛输出轨迹，对于不同信噪比输入信号，采用 SSKF 跟踪与图 7－3 给出的 EKF 跟踪结果是相同的。除了信噪比为 5dB 时发散轨迹的数目(SSKF 为 3 个，而 EKF、UKF 为 2 个)略有差异外，SSKF 与 EKF 和 UKF 在输出上没有明显差异，本书后续章节讨论均方根误差时，将会给出进一步证明。

参考文献

[1] Julier SJ. The spherical simplex unscented transformation. *Proc. Am. Control Conf*, 2003;2430－2434.

第11章 Sigma点类别:高斯-厄米特卡尔曼滤波器

一维高斯-厄米特方法最初由 Wilf[1] 首次提出,正如《数值方法》书[2]和 Ball 论文[3]中提及的那样。与前面章节提出的方法不同,高斯-厄米特积分是一种正交方法,利用正交多项式代替简单的多项式。

考虑式(5-60)中的一般多维积分,将其重写为更显式的形式,即

$$\boldsymbol{I}(\tilde{\boldsymbol{f}}) = \frac{1}{(2\pi)^{n_x/2}}\int \tilde{\boldsymbol{f}}(\boldsymbol{c})\mathrm{e}^{-\boldsymbol{c}^{\mathrm{T}}\boldsymbol{c}/2}\mathrm{d}\boldsymbol{c} \tag{11-1}$$

替换变量,消去指数中的1/2因子,令

$$\boldsymbol{q} \triangleq \frac{\boldsymbol{c}}{\sqrt{2}} = \frac{1}{\sqrt{2}}\boldsymbol{D}^{-1}(\boldsymbol{x} - \hat{\boldsymbol{x}}) \tag{11-2}$$

$$\boldsymbol{x} = \hat{\boldsymbol{x}} + \sqrt{2}\boldsymbol{D}\boldsymbol{q} \tag{11-3}$$

因此,式(11-1)变为

$$\boldsymbol{I}(\tilde{\boldsymbol{f}}) = \frac{\sqrt{2}}{(2\pi)^{n_x/2}}\int \tilde{\boldsymbol{f}}(\boldsymbol{q})\mathrm{e}^{-\boldsymbol{q}^{\mathrm{T}}\boldsymbol{q}}\mathrm{d}\boldsymbol{q} \tag{11-4}$$

$$\tilde{\boldsymbol{f}}(\boldsymbol{q}) \triangleq \boldsymbol{f}(\hat{\boldsymbol{x}} + \sqrt{2}\boldsymbol{D}\boldsymbol{q}) \tag{11-5}$$

11.1 一维高斯-厄米特求积

现在的目标是采用某种方法,将 $\tilde{f}(q)$ 展开为正交多项式。首先,考虑一维情况的积分

$$\boldsymbol{I} = \int \tilde{f}(q)\mathrm{e}^{-q^2}\mathrm{d}q \tag{11-6}$$

将 $\tilde{f}(q)$ 展开为一组正交多项式，$\tilde{f}(q)$ 可以用下式逼近

$$\tilde{f}(q) \simeq \sum_{l=0}^{m} p_l \phi_l(q) \tag{11-7}$$

这里 $\{\phi_l(q), l=0,1,\cdots,m\}$ 代表一组正交多项式，多项式 $\phi_l(q)$ 和 $\phi_k(q)$ 与权重函数（或核函数）$p(q)$ 在区间 $[a,b]$ 上正交，即

$$\int_a^b p(q)\phi_l(q)\phi_k(q)\mathrm{d}q = c_l \delta_l k \tag{11-8}$$

这里 c_l 为归一化常量，而

$$\delta_{lk} = \begin{cases} 1 & k = l \\ 0 & k \neq l \end{cases} \tag{11-9}$$

对于式(11-6)的积分，可以使用厄米特多项式 $H_l(q)$，权值函数为

$$p(q) = \mathrm{e}^{-q^2} \tag{11-10}$$

积分区间为 $[a,b] \to [-\infty, \infty]$，式(11-8)正交条件则变为

$$\int_{-\infty}^{\infty} \mathrm{e}^{-q^2} H_l(q) H_k(q)\mathrm{d}q = h_l \delta_{lk} \tag{11-11}$$

这里，归一化常量由下式给出，即

$$h_l = \sqrt{\pi} 2^l l! \tag{11-12}$$

可以利用递归关系[3]生成厄米特多项式

$$H_{-1}(q) = 0 \tag{11-13}$$

$$H_0(q) = 1 \tag{11-14}$$

$$H_{l+1}(q) = 2qH_l(q) - 2lH_{l-1}(q) \quad l \geqslant 0 \tag{11-15}$$

将规范化的厄米特多项式定义为

$$\widetilde{H}_l(q) \triangleq \frac{H_l(q)}{\sqrt{h_l}} \tag{11-16}$$

进而导出正交条件

$$\int_{-\infty}^{\infty} \mathrm{e}^{-q^2} \widetilde{H}_l(q) \widetilde{H}_k(q)\mathrm{d}q = \delta_{lk} \tag{11-17}$$

正交厄米特多项式的递归关系为

$$\widetilde{H}_{-1}(q) = 0 \tag{11-18}$$

$$\widetilde{H}_0(q) = h_0^{-1/2} = \pi^{-1/4} \tag{11-19}$$

$$\widetilde{H}_{l+1}(q) = q\sqrt{\frac{2}{l+1}}\widetilde{H}_l(q) - \sqrt{\frac{l}{l+1}}\widetilde{H}_{l-1}(q) \quad l \geqslant 0 \tag{11-20}$$

如果定义

$$\beta_l \triangleq \sqrt{l/2} \tag{11-21}$$

那么式(11－20)变为

$$q\widetilde{H}_l(q) = \beta_l\widetilde{H}_{l-1}(q) + \beta_{l+1}\widetilde{H}_{l+1}(q) \tag{11-22}$$

将矢量 $\boldsymbol{h}(q)$、$\boldsymbol{e}_m$ 以及矩阵 $\boldsymbol{J}_m$ 定义为

$$\boldsymbol{h}(q) \triangleq [\widetilde{H}_0(q), \widetilde{H}_1(q), \cdots, \widetilde{H}_{m-1}(q)]^{\mathrm{T}} \tag{11-23}$$

$$\boldsymbol{e}_m \triangleq [0, 0, \cdots, 0, 1]^{\mathrm{T}} \tag{11-24}$$

$$\boldsymbol{J}_m \triangleq \begin{bmatrix} 0 & \beta_1 & \cdots & 0 \\ \beta_1 & 0 & & \vdots \\ \vdots & & \ddots & \beta_{m-1} \\ 0 & \cdots & \beta_{m-1} & 0 \end{bmatrix} \tag{11-25}$$

则式(11－22)的正交厄米特递推关系可重写为

$$q\boldsymbol{h}(q) = \boldsymbol{J}_m\boldsymbol{h}(q) + \beta_m\widetilde{H}_m(q)\boldsymbol{e}_l \tag{11-26}$$

设

$$\widetilde{H}_m(q) = 0 \tag{11-27}$$

则式(11－26)变为特征值方程,即

$$[\boldsymbol{J}_m - q\boldsymbol{I}]\boldsymbol{h}(q) = \boldsymbol{0} \tag{11-28}$$

$\boldsymbol{J}_m$ 的 m 个不同特征值由$\{q_l, l=0,1,\cdots,m-1\}$给出,为 $\tilde{f}(q)$的 m 阶展开式的插值点。$\tilde{f}(q)$在不同的特征值$\{q_l, l=0,1,\cdots,m-1\}$处插值误差为零[2],由于式(11－17)正交条件中包括权值函数,特征值可以写成式(11－27)归一化解的集合,即

$$q_l = \frac{\widetilde{H}_m(ql)}{p(q_l)} \quad l = 0,1,\cdots,m-1 \tag{11-29}$$

这里 $p(q_l)$由式(11－10)给出。

考虑 $m=3$ 时三阶多项式的展开。尽管高阶多项式展开方法是由一维情况

推广而来(如文献[4,5]),但是当将高阶方法应用到多维情况时,很快就变得难以控制。对于 $m=3$,式(11-25)变为

$$\boldsymbol{J}_3=\begin{bmatrix}0&\beta_1&0\\\beta_1&0&\beta_2\\0&\beta_2&0\end{bmatrix}=\begin{bmatrix}0&1/\sqrt{2}&0\\1/\sqrt{2}&0&1\\0&1&0\end{bmatrix}\tag{11-30}$$

式(11-30)的特征值由如下集合构成,即

$$\{q_0,q_1,q_2\}=\frac{\sqrt{3}}{2}\{-1,0,1\}\tag{11-31}$$

令 $\boldsymbol{r}\triangleq\{0,1,-1\}$,式(11-31)可重写为

$$\{q_0,q_1,q_2\}=\frac{\sqrt{3}}{2}\boldsymbol{r}\tag{11-32}$$

函数 $\tilde{f}(q)$ 现在可以用规范化厄米特多项式的三阶展开式逼近,即

$$\tilde{f}(q)\simeq\sum_{l=0}^{2}a_l\widetilde{H}_l(q)\tag{11-33}$$

由于 $\tilde{f}(q)$ 在特征值点的展开是准确的,将每个特征点代入式(11-33),得矩方程组为

$$\tilde{f}(q_0)\simeq\sum_{l=0}^{2}a_l\widetilde{H}_l(q_0)\tag{11-34}$$

$$\tilde{f}(q_1)\simeq\sum_{l=0}^{2}a_l\widetilde{H}_l(q_1)\tag{11-35}$$

$$\tilde{f}(q_2)\simeq\sum_{l=0}^{2}a_l\widetilde{H}_l(q_2)\tag{11-36}$$

这三个方程有三个未知数 $\{a_1,a_2,a_3\}$,未知系数是可解的。将矢量 $\tilde{\boldsymbol{f}}$、$\boldsymbol{a}$ 以及矩阵 $\widetilde{\boldsymbol{H}}$ 定义为

$$\tilde{\boldsymbol{f}}\triangleq[\tilde{f}(q_0),\tilde{f}(q_1),\tilde{f}(q_2)]^{\mathrm{T}}\tag{11-37}$$

$$\boldsymbol{a}\triangleq[a_0,a_1,a_2]^{\mathrm{T}}\tag{11-38}$$

$$\widetilde{\boldsymbol{H}}\triangleq\begin{bmatrix}\widetilde{H}_0(q_0)&\widetilde{H}_1(q_0)&\widetilde{H}_2(q_0)\\\widetilde{H}_0(q_1)&\widetilde{H}_1(q_1)&\widetilde{H}_2(q_1)\\\widetilde{H}_0(q_2)&\widetilde{H}_1(q_2)&\widetilde{H}_2(q_2)\end{bmatrix}\tag{11-39}$$

方程组(11-34)~式(11-36)可重写矢量矩阵方程为

$$\tilde{\boldsymbol{f}}=\widetilde{\boldsymbol{H}}\boldsymbol{a}\rightarrow\boldsymbol{a}=\widetilde{\boldsymbol{H}}^{-1}\tilde{\boldsymbol{f}}\tag{11-40}$$

将式(11-12)~式(11-16)代入式(11-32),得

$$\widetilde{\boldsymbol{H}} = \frac{1}{\pi^{1/4}}\begin{bmatrix} 1 & 0 & -1/\sqrt{2} \\ 1 & \sqrt{3} & \sqrt{2} \\ 1 & -\sqrt{3} & \sqrt{2} \end{bmatrix} \tag{11-41}$$

$\widetilde{\boldsymbol{H}}$ 的逆为

$$\widetilde{\boldsymbol{H}}^{-1} = \pi^{1/4}\begin{bmatrix} 2/3 & 1/6 & 1/6 \\ 0 & 1/2\sqrt{3} & -1/2\sqrt{3} \\ -\sqrt{2}/3 & 1/3\sqrt{2} & 1/3\sqrt{2} \end{bmatrix} \tag{11-42}$$

将式(11－42)代入式(11－40)，得

$$\begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \pi^{1/4}\begin{bmatrix} 2/3 & 1/6 & 1/6 \\ 0 & 1/2\sqrt{3} & -1/2\sqrt{3} \\ -\sqrt{2}/3 & 1/3\sqrt{2} & 1/3\sqrt{2} \end{bmatrix}\begin{bmatrix} \tilde{f}(0) \\ \tilde{f}\left(\frac{\sqrt{3}}{2}\right) \\ \tilde{f}\left(\frac{-\sqrt{3}}{2}\right) \end{bmatrix} \tag{11-43}$$

求解式(11－43)，得$\{a_0, a_1, a_2\}$为

$$a_0 = \pi^{1/4}\left[\frac{2}{3}\tilde{f}(0) + \frac{1}{6}\tilde{f}\left(\sqrt{\frac{3}{2}}\right) + \frac{1}{6}\tilde{f}\left(-\sqrt{\frac{3}{2}}\right)\right] \tag{11-44}$$

$$a_1 = \frac{\pi^{1/4}}{2\sqrt{3}}\left[\tilde{f}\left(\sqrt{\frac{3}{2}}\right) - \tilde{f}\left(-\sqrt{\frac{3}{2}}\right)\right] \tag{11-45}$$

$$a_2 = \frac{\pi^{1/4}}{3\sqrt{2}}\left[\tilde{f}\left(\sqrt{\frac{3}{2}}\right) - 2\tilde{f}(0) + \tilde{f}\left(-\sqrt{\frac{3}{2}}\right)\right] \tag{11-46}$$

在 $\tilde{f}(q)$ 的厄米特展开式中，仅保留第一项，则式(11－33)变成

$$\tilde{f}(q) = a_0\widetilde{H}_0(q) \tag{11-47}$$

$$\tilde{f}(q) = \frac{2}{3}\tilde{f}(0) + \frac{1}{6}\tilde{f}\left(\sqrt{\frac{3}{2}}\right) + \frac{1}{6}\tilde{f}\left(-\sqrt{\frac{3}{2}}\right) \tag{11-48}$$

为避免混淆，在多维情况时，将一维插值点集合及其权值定义为

$$\{q^{(0)}, q^{(1)}, q^{(2)}\} = \left\{0, \sqrt{\frac{3}{2}}, -\sqrt{\frac{3}{2}}\right\} \tag{11-49}$$

$$\{p_0, p_1, p_2\} = \{2/3, 1/6, 1/6\} \tag{11-50}$$

这样，式(11－48)的厄米特多项展开式变为

$$\tilde{f}(q) \simeq \sum_{i=0}^{2} p_i\,\tilde{f}(q^{(i)}) \tag{11-51}$$

式中:$q^{(i)}$ 和 p_i 分别由式(11－49)和式(11－50)给出。

11.2 一维高斯－厄米特卡尔曼滤波器

一维情况下,式(11－4)变为

$$I(\tilde{f}) = \frac{1}{\sqrt{\pi}}\int \tilde{f}(q)\mathrm{e}^{-q^2}\mathrm{d}q \tag{11-52}$$

$$q = \frac{c}{\sqrt{2}} = \frac{1}{\sqrt{2}}\sigma^{-1}(x - \hat{x}) \tag{11-53}$$

$$x = \hat{x} + \sqrt{2}\sigma q \tag{11-54}$$

将式(11－49)代入式(11－54),得一维高斯厄米特 Sigma 点为

$$\chi^{(i)} = \hat{x} + \sqrt{2}\sigma q^{(i)} \tag{11-55}$$

此外,有

$$\tilde{f}(q^{(i)}) = \tilde{f}(\hat{x} + \sqrt{2}\sigma q^{(i)}) = f(\chi^{(i)}) \tag{11-56}$$

则式(11－52)变为

$$\begin{aligned} I(\tilde{f}) &= \sum_{i=0}^{2} p_i f(\chi^{(i)}) \frac{1}{\sqrt{\pi}}\int \mathrm{e}^{-q^2}\mathrm{d}q = \sum_{i=0}^{2} p_i f(\chi^{(i)}) \frac{1}{\sqrt{2\pi}}\int \mathrm{e}^{-c^2/2}\mathrm{d}c \\ &= \sum_{i=0}^{2} p_i f(\chi^{(i)}) \int \mathcal{N}(c;0,1)\mathrm{d}c \end{aligned} \tag{11-57}$$

由于密度函数的积分等于 1,式(11－57)可简化为

$$I(\tilde{f}) = \sum_{i=0}^{2} p_i f(\chi^{(i)}) \tag{11-58}$$

利用式(11－58),一维状态预测方程(5－51)可写为

$$\hat{x}_{n|n-1} = \sum_{i=0}^{2} p_i f(\chi_{n-1|n-1}^{(i)}) \tag{11-59}$$

$$\chi_{n-1|n-1}^{(i)} = \hat{x}_{n-1|n-1} + \sqrt{3}\sigma_{xx,n-1|n-1}\boldsymbol{r}^{(i)} \tag{11-60}$$

$$\boldsymbol{r} \triangleq [1] \in \mathbf{R}^1 = \begin{cases} 0 & i = 0 \\ 1 & i = 1 \\ -1 & i = 2 \end{cases} \tag{11-61}$$

对于一维状态变量预测,令

$$\tilde{g}(c) \triangleq [\tilde{f}(c) - \hat{x}_{n|n-1}]^2 \tag{11-62}$$

则一维状态变量预测方程(7-2)变成

$$\sigma_{xx,n|n-1}^{2} = \int \tilde{g}(c)\ \mathcal{N}(c;0,1)\,\mathrm{d}c + \sigma_{v,n}^{2} \tag{11-63}$$

注意:式(11-63)中的积分项与式(11-6)相同,所以有

$$\sigma_{xx,n|n-1}^{2} = \sum_{i=0}^{2} p_i[\tilde{f}(\chi_{n-1|n-1}^{(i)}) - \hat{x}_{n|n-1}]^2 + \sigma_{v,n}^{2} \tag{11-64}$$

用同样的方式,将式(5-51)~式(5-52)简化为

$$\hat{z}_{n|n-1} = \sum_{i=0}^{2} p_i h(\chi_{n|n-1}^{(i)}) \tag{11-65}$$

$$\sigma_{zz,n|n-1}^{2} = \sum_{i=0}^{2} p_i[f(\chi_{n|n-1}^{(i)}) - \hat{x}_{n|n-1}]^2 + \sigma_{w,n}^{2} \tag{11-66}$$

$$\sigma_{xz,n|n-1}^{2} = \sum_{i=0}^{2} p_i[f(\chi_{n|n-1}^{(i)}) - \hat{x}_{n|n-1}]\,[h(\chi_{n|n-1}^{(i)}) - \hat{z}_{n|n-1}] \tag{11-67}$$

$$\chi_{n|n-1}^{(i)} = \hat{x}_{n|n-1} + \sqrt{3}\sigma_{xx,n|n-1}\boldsymbol{r}^{(i)} \tag{11-68}$$

式(11-59)和式(11-64)~式(11-67)构成了一维高斯-厄米特卡尔曼滤波器预测方程组。这里没有给出一维情况的过程流图,不过可以把它作为下节多维情况的一个特例。

11.3 多维高斯-厄米特卡尔曼滤波器

如果要先构建插值多维多项式,那么构建多维正交多项式积分法则是一个很难实现的任务。构建多维容积规则的最简单方法是由一维积分法则构建乘积法则。首先考虑二维情况,$\boldsymbol{q} = [q_1, q_2]^{\mathrm{T}}$,将 $\tilde{f}(\boldsymbol{q})$ 展开为二重累加和,即

$$\tilde{\boldsymbol{f}}(\boldsymbol{q}) = \sum_{i_1=0}^{2}\sum_{i_2=0}^{2} p_{i_1}p_{i_2}\tilde{\boldsymbol{f}}(q_1^{(i_1)}, q_2^{(i_2)}) = \sum_{j=0}^{3^2-1} w_j \tilde{\boldsymbol{f}}(\boldsymbol{q}^{(j)}) \tag{11-69}$$

式中:p_{i_1} 和 p_{i_2} 由式(11-50)定义;$q_1^{(i_1)}$ 和 $q_2^{(i_2)}$ 由式(11-49)定义。式(11-69)的双重累加和可以变换成索引为$\{j=0,1,\cdots,8\}$的单重累加和,权值变为

$$w_0 = \left(\frac{2}{3}\right)^2 = p_0^2 \tag{11-70}$$

$$w_1 = w_2 = w_3 = w_4 = \left(\frac{2}{3}\right)\left(\frac{1}{6}\right) = p_0 p_i = p_i p_0 \quad i = 1,2 \tag{11-71}$$

$$w_5 = w_6 = w_7 = w_8 = \left(\frac{1}{6}\right)^2 = p_i p_l = p_l p_i \quad i,l = 1,2 \tag{11-72}$$

将式(11-70)~式(11-72)可写为更紧凑的形式,即

$$w_j \triangleq \begin{cases} \left(\frac{2}{3}\right)^2 & j = 0 \\ \left(\frac{2}{3}\right)\left(\frac{1}{6}\right) & j = 1, \cdots, 4 \\ \left(\frac{1}{6}\right)^2 & j = 5, \cdots, 8 \end{cases} \tag{11-73}$$

由式(11－50),得

$$\boldsymbol{q}^{(j)} = \sqrt{\frac{3}{2}}\boldsymbol{r}^{(j)} \tag{11-74}$$

$$\boldsymbol{r}^{(j)} = \begin{cases} [0] \in \mathbf{R}^2 = [0,0]^{\mathrm{T}} & j = 0 \\ [1] \in \mathbf{R}^2 = \begin{cases} [1,0]^{\mathrm{T}} & j = 1 \\ [0,1]^{\mathrm{T}} & j = 2 \\ [-1,0]^{\mathrm{T}} & j = 3 \\ [0,-1]^{\mathrm{T}} & j = 4 \end{cases} \\ [1,1] \in \mathbf{R}^2 = \begin{cases} [1,1]^{\mathrm{T}} & j = 5 \\ [-1,1]^{\mathrm{T}} & j = 6 \\ [1,-1]^{\mathrm{T}} & j = 7 \\ [-1,-1]^{\mathrm{T}} & j = 8 \end{cases} \end{cases} \tag{11-75}$$

对于4维状态矢量采用同样的方式,$\tilde{\boldsymbol{f}}(\boldsymbol{q})$展开式可写为

$$\begin{aligned} \tilde{\boldsymbol{f}}(\boldsymbol{q}) &= \sum_{i_1=0}^{2}\sum_{i_2=0}^{2}\sum_{i_3=0}^{2}\sum_{i_4=0}^{2} p_{i_1}p_{i_2}p_{i_3}p_{i_4}\tilde{\boldsymbol{f}}(q_1^{(i_1)}, q_2^{(i_2)}, q_1^{(i_3)}, q_2^{(i_4)}) \\ &= \sum_{j=0}^{3^4-1} w_j \tilde{\boldsymbol{f}}(\boldsymbol{q}^{(j)}) \end{aligned} \tag{11-76}$$

采用与推导式(11－73)相同的p的乘积组合方式,得权值集合

$$w_j = \begin{cases} \left(\frac{2}{3}\right)^4 & j = 0 \\ \left(\frac{2}{3}\right)^3\left(\frac{1}{6}\right) & j = 1, \cdots, 8 \\ \left(\frac{2}{3}\right)^2\left(\frac{1}{6}\right)^2 & j = 9, \cdots, 32 \\ \left(\frac{2}{3}\right)\left(\frac{1}{6}\right)^3 & j = 33, \cdots, 64 \\ \left(\frac{1}{6}\right)^4 & j = 65, \cdots, 80 \end{cases} \tag{11-77}$$

由式(11－50),得

$$\boldsymbol{q}^{(j)} = \sqrt{\frac{3}{2}}\boldsymbol{r}^{(j)} \tag{11-78}$$

$$\boldsymbol{r}^{(j)} = \begin{cases} [0] \in \mathbf{R}^4 & j = 0 \\ [1] \in \mathbf{R}^4 & j = 1,\cdots,8 \\ [1,1] \in \mathbf{R}^4 & j = 9,\cdots,32 \\ [1,1,1] \in \mathbf{R}^4 & j = 33,\cdots,64 \\ [1,1,1,1] \in \mathbf{R}^4 & j = 65,\cdots,80 \end{cases} \tag{11-79}$$

这可以推广至一般的 n_x 维情况,有

$$\tilde{\boldsymbol{f}}(\boldsymbol{q}) = \sum_{j=0}^{3^{n_x-1}} w_j \tilde{\boldsymbol{f}}(\boldsymbol{q}^{(j)}) \tag{11-80}$$

这里 $\boldsymbol{q}^{(j)}$ 由式(11-78)给出,且有

$$\boldsymbol{r}^{(j)} = \begin{cases} [0] \in \mathbf{R}^n & j = 0 \\ [1] \in \mathbf{R}^n & j = 1,\cdots,2n \\ [1,1] \in \mathbf{R}^n & j = 2n+1,\cdots,2(n+n^{(2)}) \\ [1,1,1] \in \mathbf{R}^n & j = [2(n+n^{(2)}+1],\cdots, \\ & \quad [2(n+n^{(2)})+2^3 n_{(3)}] \\ \vdots & \\ [1,1,\cdots,1] \in \mathbf{R}^n & j = [2(n+n^{(2)}) + \sum_{i=3}^{n-2} 2^i n_{(i)}] \\ & \quad + [1,\cdots,2^{n-1} n_{(n-1)}] \\ [1,1,1,\cdots,1] \in \mathbf{R}^n & j = [2(n+n^{(2)}) + \sum_{i=3}^{n-1} 2^i n_{(i)}] \\ & \quad + [1,\cdots,2^n n_{(n)}] \end{cases} \tag{11-81}$$

为方便起见,用 n 代替 n_x,并分别用式(2-24)和式(2-25)对 $n^{(i)}$ 和 $n_{(i)}$ 进行定义。对于最后的发生器,有 $n \pm 1$ 个重复。权值可表示为

$$w_j = \begin{cases} \left(\frac{2}{3}\right)^n & j = 0 \\ \left(\frac{2}{3}\right)^{n-1}\left(\frac{1}{6}\right) & j = 1,\cdots,2n \\ \left(\frac{2}{3}\right)^{n-2}\left(\frac{1}{6}\right)^2 & j = 2n+1,\cdots,2(n+n^{(2)}) \\ \vdots & \\ \left(\frac{2}{3}\right)\left(\frac{1}{6}\right)^{n-1} & j = [2(n+n^{(2)}) + \sum_{i=3}^{n-2} 2^i n_{(i)}] \\ & \quad + [1,\cdots,2^{n-1} n_{(n-1)}] \\ \left(\frac{1}{6}\right)^n & j = [2(n+n^{(2)}) + \sum_{i=3}^{n-1} 2^i n_{(i)}] \\ & \quad + [1,\cdots,2^n n_{(n)}] \end{cases} \tag{11-82}$$

现在,将式(11-80)给出的 $\tilde{f}(\boldsymbol{q})$ 的一般厄米特多项展开式代入矩积分方程(11-4),得

$$\boldsymbol{I}(\tilde{\boldsymbol{f}}) = \frac{\sqrt{2}}{(2\pi)^{n_x/2}}\int \sum_{j=0}^{3^{n_x}-1} w_j \tilde{\boldsymbol{f}}(\boldsymbol{q}^{(j)}) \mathrm{e}^{-\boldsymbol{q}^{\mathrm{T}}\boldsymbol{q}} \mathrm{d}\boldsymbol{q} \tag{11-83}$$

将变量变回 $\boldsymbol{c}$ 的积分,式(11-83)变成

$$\boldsymbol{I}(\tilde{\boldsymbol{f}}) = \sum_{j=0}^{3^{n_x}-1} w_j \tilde{\boldsymbol{f}}(\boldsymbol{q}^{(j)}) \int \mathcal{N}(\boldsymbol{c};\boldsymbol{0},\boldsymbol{I}) \mathrm{d}\boldsymbol{c} = \sum_{j=0}^{3^{n_x}-1} w_j \tilde{\boldsymbol{f}}(\boldsymbol{q}^{(j)}) \tag{11-84}$$

利用式(11-5)变换 $\tilde{f}(\boldsymbol{q}^{(j)})$,则式(11-84)变成

$$\boldsymbol{I} = \sum_{j=0}^{3^{n_x}-1} w_j \tilde{\boldsymbol{f}}(\hat{\boldsymbol{x}} + \sqrt{3}\boldsymbol{D}\boldsymbol{r}^{(j)}) \tag{11-85}$$

令

$$\boldsymbol{\chi}_{n-1|n-1}^{(j)} \triangleq \hat{\boldsymbol{x}}_{n-1|n-1} + \sqrt{3}\boldsymbol{D}_{n-1|n-1}\boldsymbol{r}^{(j)} \tag{11-86}$$

利用式(11-85)计算式(5-51)~式(5-52),得一般高斯-厄米特状态预测方程为

$$\hat{\boldsymbol{x}}_{n|n-1} = \sum_{j=0}^{3^{n_x}-1} w_j \boldsymbol{f}(\boldsymbol{\chi}_{n-1|n-1}^{(j)}) \tag{11-87}$$

$$\boldsymbol{P}_{n|n-1}^{xx} = \sum_{j=0}^{3^{n_x}-1} w_j [\boldsymbol{f}(\boldsymbol{\chi}_{n-1|n-1}^{(j)}) - \boldsymbol{x}_{n|n-1}] \cdot [\boldsymbol{f}(\boldsymbol{\chi}_{n-1|n-1}^{(j)}) - \hat{\boldsymbol{x}}_{n|n-1}]^{\mathrm{T}} + \boldsymbol{Q} \tag{11-88}$$

同样,定义

$$\boldsymbol{\chi}_{n|n-1}^{(j)} \triangleq \hat{\boldsymbol{x}}_{n|n-1} + \sqrt{3}\boldsymbol{D}_{n|n-1}\boldsymbol{r}^{(j)} \tag{11-89}$$

观测值预测方程式(5-55)~式(5-57)可写为

$$\hat{\boldsymbol{z}}_{n|n-1} = \sum_{j=0}^{3^{n_x}-1} w_j \boldsymbol{h}(\boldsymbol{\chi}_{n|n-1}^{(j)}) \tag{11-90}$$

$$\boldsymbol{P}_{n|n-1}^{zz} = \sum_{j=0}^{3^{n_x}-1} w_j [\boldsymbol{h}(\boldsymbol{\chi}_{n|n-1}^{(j)}) - \hat{\boldsymbol{x}}_{n|n-1}] \cdot \tag{11-91}$$

$$[\boldsymbol{h}(\boldsymbol{\chi}_{n|n-1}^{(j)}) - \hat{\boldsymbol{x}}_{n|n-1}]^{\mathrm{T}} + \boldsymbol{R} \tag{11-92}$$

$$\boldsymbol{P}_{n|n-1}^{xz} = \sum_{j=0}^{3^{n_x}-1} w_j [\tilde{\boldsymbol{f}}(\boldsymbol{\chi}_{n|n-1}^{(j)}) - \hat{\boldsymbol{x}}_{n|n-1}] \cdot \tag{11-93}$$

$$[\boldsymbol{h}(\boldsymbol{\chi}_{n|n-1}^{(j)}) - \hat{\boldsymbol{x}}_{n|n-1}]^{\mathrm{T}} \tag{11-94}$$

注意:高斯-厄米特预测方程与 UKF 和 SSKF 的预测方程相同,除了求和的上限外,与 FDKF 的第二种形式也相同。第 13 章将总结这些 Sigma 点卡尔曼滤波器,并且给出过程框图。表 11-1 概括了 GHKF 过程。

表 11－1　多维高斯－厄米特卡尔曼滤波器过程

步骤(1)	滤波器初始化	初始化和 $\hat{x}_0$ 和 P_0^{xx}
步骤(2)	状态矢量预测	$\chi_{n-1\|n-1}^{(j)} = \hat{x}_{n-1\|n-1} + \sqrt{3}D_{n-1\|n-1}r^{(j)} \quad j=0,1,\cdots,3^{n_x}-1$ $\hat{x}_{n\|n-1} = \sum_{j=0}^{3n_x-1} w_j f(\chi_{n-1\|n-1}^{(j)})$ $P_{n\|n-1}^{xx} = \sum_{j=0}^{3n_x-1} w_j[f(\chi_{n-1\|n-1}^{(j)}) - \hat{x}_{n\|n-1}] \cdot [f(\chi_{n-1\|n-1}^{(j)}) - \hat{x}_{n\|n-1}]^{\mathrm{T}} + Q$
步骤(3)	观测值预测	$\chi_{n\|n-1}^{(j)} = \hat{x}_{n\|n-1} + \sqrt{3}D_{n\|n-1}r^{(j)} \quad j=0,1,\cdots,3^{n_x}-1$ $\hat{z}_{n\|n-1} = \sum_{j=0}^{3n_x-1} w_j \tilde{h}(\chi_{n\|n-1}^{(j)})$ $P_{n\|n-1}^{zz} = \sum_{j=0}^{3n_x-1} w_j[h(\chi_{n\|n-1}^{(j)}) - \hat{z}_{n\|n-1}][h(\chi_{n\|n-1}^{(j)}) - \hat{z}_{n\|n-1}]^{\mathrm{T}} + R$ $P_{n\|n-1}^{xz} = \sum_{j=0}^{3n_x-1} w_j[f(\chi_{n-1\|n-1}^{(j)}) - \hat{x}_{n\|n-1}] \cdot [h(\chi_{n\|n-1}^{(j)}) - \hat{z}_{n\|n-1}]^{\mathrm{T}}$
步骤(4)	卡尔曼滤波器更新	$K_n \triangleq P_{n\|n-1}^{xz}(P_{n\|n-1}^{zz})^{-1}$ $\hat{x}_{n\|n} = \hat{x}_{n\|n-1} + K_n(z_n^o - \hat{z}_{n\|n-1})$ $P_{n\|n}^{xx} = P_{n\|n-1}^{xx} - K_n P_{n\|n-1}^{zz} K_n^{\mathrm{T}}$
步骤(5)	存储结果	将 $\hat{x}_{n\|n}$ 和 $P_{n\|n}^{xx}$ 存储到跟踪文件中
步骤(6)	返回步骤(2)	

11.4　高维/高阶多项式的稀疏网格逼近

如上所述，高斯－厄米特滤波器所需网格数目为

$$N_{gh} = m^{n_x} \tag{11-95}$$

式中：m 为用于逼近的厄米特多项式的阶数；n_x 为状态矢量的维度。当状态矢量的维数增加时，所需 Sigma 点的数量呈指数级增长，导致“维数灾难”；当维数大于4 时，完全填充的方形网格的 Sigma 点数目变得非常大，此时很少使用 GHKF滤波器。本章前面几节，已经给出二阶（$m=3$）厄米特展开需要 3^{n_x} 个 Sigma点，对于这种情况，减少点的数目的唯一方法是用 UKF 代替 GHKF，不过这样会损失一些精度。

在式（11－30）前述内容可知，厄米特插值方法可以推广至高阶。当然，高阶插值需要更多的 Sigma 点，使维数灾难的情况更为严重。由于高阶厄米特插值需要求解与厄米特多项式同阶的多项式方程，通常采用数值方法来生成特征值，得权值和 Sigma 点，用无理数替代式（11－48）中的有理数。文献[7]给出了这些无理数的列表。然而，当使用数据列表时必须注意，权值是为非规范化的厄米特多项式而生成的。

多维正交积分的最新进展 - Smolyak[8]的工作,使得高阶高斯 - 厄米特积分更易于实现。Smolyak 提出一种多维正交积分的数值方法,以牺牲准确性为代价,减少了网格点的数量。完整的减少网格点的数学方法介绍可参考文献[9 - 13]。上面介绍的厄米特 Sigma 点的生成方法符合 Smolyak 的方法,这里 Sigma 点根据它们与原点的距离,按照层次进行分组,在式(11 - 81)一般维度情况下可以看到。对稀疏网格数值积分,必须取整个网格的一个子集,并重新归一化缩减的权值集合。可惜,这种稀疏网格积分方法不能减少厄米特二阶多项式所需要的Sigma点的数量。

例如,仅需要两个最低级的网格点,$\boldsymbol{r}^{(0)} = [0] \in \mathbf{R}^{n_x}, j=0$ 和 $\boldsymbol{r}^{(j)} = [1] \in \mathbf{R}^{n_x}, j=1,\cdots,n_x$。初始权值为 $\omega_0 = (2/3)^{n_x}$ 和 $\omega_j = (2/3)^{n_x-1}(1/6), j=1,\cdots,n_x$。每个权值必须除以所有的权重的总和$(2/3)^{n_x-1}([4+n_x]/6)$,实现归一化。归一化后的权值为 $\omega_0 = 4/(4+n_n)$ 和 $\omega_j = 1/(4+n_n), j=1,\cdots,n_x$。这样稀疏高斯 - 厄米特滤波器与式(9 - 24)、式(9 - 25)的 UKF 在形式上相同,但是有

$$\boldsymbol{c}^{(j)} = q\boldsymbol{r}^{(j)} = \sqrt{\frac{3}{2}}\boldsymbol{r}^{(j)} \quad j=0,1,\cdots,2n_x \tag{11-96}$$

$$w_j = \begin{cases} 4/(4+n) & j=0 \\ \dfrac{1}{4}w_0 & j=1,\cdots,2n_x \end{cases} \tag{11-97}$$

以类似的方式,可以从式(11 - 81)中得尽可能多的 Sigma 网格点的子集,并且在将权值归一化后,生成另一种稀疏网格 GHKF 滤波器。该方法使得稀疏网格 GHKF 适用于高维情况。此外,对于高度非线性情况,可使用高阶厄米特多项式来得高阶稀疏网格 GHKF 滤波器,以合理的计算成本提供更高的精度。Jia 等[14]首次将这一方法应用到航天器姿态估计。在他们的论文中给出了不同稀疏度 GHKF 滤波器所需点的数目的列表,其中厄米特多项式达到 7 阶、状态矢量达到 6 维。

11.5 GHKF 在 DIFAR 船舶跟踪的应用案例分析

对于 DIFAR 跟踪问题,动态方程是线性的,处理这一问题用到的 GHKF 过程与表 9 - 2 所列完全相同。对于权值和 Sigma 点,$\{(\omega_j, \boldsymbol{c}^{(j)}), j=0,1,\cdots,n_x\}$ 的值由式(11 - 77) ~ 式(11 - 79)给出,这里取 $\boldsymbol{c}^{(j)} = \sqrt{3}\boldsymbol{r}^{(j)}$。在这个 GHKF 滤波器中,采用了与 EKF 相同的初始化以及动态噪声值 q(分别在 7.4.3 节和 7.4.4 节中)。比较 100 次蒙特卡洛输出轨迹,结果表明,在不同信噪比条件下,GHKF 滤波器的输出结果与 EKF 的输出几乎没有区别,甚至连 GHKF 发散航迹的数目与 EKF 蒙特卡洛跟踪的数目也相同,所以这里不再给出 GHKF 的跟踪结果比较图。

参考文献

[1] Wilf HS. *Mathematics for the Physical Sciences.* New York: Wiley, 1962: 54.

[2] Press WH, et. al. *Numerical Recipes in C.* Cambridge, UK: Cambridge University Press, 1992.

[3] Ball JS. Orthogonal polynomials, Gaussian quadratures and PDEs. *Comput. Sci. & Eng*, 1999: 92 – 95.

[4] Liu Q, Jackel P. A Note on multivariate Gauss – Hermite quadrature. *Biometrika*, 1994, 81(3): 624 – 629.

[5] Nia VP. Gauss – Hermite quadrature: *Numer. Stat. Method* 2006, Online: vahid. probstat. ch/paper/ ghq. pdf.

[6] Arasaratnam I, Haykin S, Elliot RJ. Discrete – time nonlinear filtering algorithms using Gauss – Hermite quadrature. *Proc. IEEE*, 2007, 95(5): 953 – 977.

[7] Abramowitz M, Stegun IA. *Handbook of Mathematical Functions.* Dover Publications, 1972.

[8] Smolyak SA. Quadrature and interpolation formulas for tensor products of certain classes of functions. *Dokl. Akad. Nauk SSSR*, 1963, 4: 240 – 243.

[9] Wasilkowski GW, Wozniakowski H. Explicit cost bounds of algorithms for multivariate tensor product problems. *J. Complex*, 1994, 11(1): 1 – 56.

[10] Gerstner T, Griebel M. Numerical integration using sparse grids. *Numer. Algorithms*, 1998, 18(4): 209 – 232.

[11] Gerstner T, Griebel M. Dimension – adaptive tensor – product quadrature. *Computing*, 2003, 71: 65 – 87

[12] Heiss F, Winschel V. Estimation with Numerical Integration on Sparse Grids. Discussion Papers in Economics 916, 2006.

[13] Heiss F, Winschel V. Likelihood approximation by numerical integration on sparse grids. J. Econ, 2008, 144: 62 – 80.

[14] Jia B, Xin M, Cheng Y. Sparse Gauss – Hermite quadrature filter for spacecraft attitude estimation. Am. Control Conf, 2010, 15(4): 2873 – 2878.

第12章 蒙特卡洛卡尔曼滤波器

Sigma 点卡尔曼滤波器首先对状态矢量进行仿射变换，然后将非线性函数展开为多项式，将高斯加权积分简化为矩方程组，通过求解方程组，得到一组权值和 Sigma 点。虽然 Sigma 点方法在一些文献中被用来逼近密度函数，但在本书的推导中并没用其进行高斯密度的逼近。

这里介绍一种通过一组蒙特卡罗采样逼近高斯密度的方法。回顾式(5-51)，得

$$\hat{\boldsymbol{x}}_{n|n-1} = \int \tilde{\boldsymbol{f}}(\boldsymbol{c})\ \mathcal{N}(\boldsymbol{c};\boldsymbol{0},\boldsymbol{I})\mathrm{d}\boldsymbol{c} \tag{12-1}$$

假设由$\mathcal{N}(\boldsymbol{c};\boldsymbol{0},\boldsymbol{I})$生成一组蒙特卡洛斯采样$\{\boldsymbol{c}^{(j)},j=1,\cdots,N_s\}$。将$\mathcal{N}(\boldsymbol{c};\boldsymbol{0},\boldsymbol{I})$表示为密度函数的离散逼近，即

$$\mathcal{N}(\boldsymbol{c};\boldsymbol{0},\boldsymbol{I}) \simeq \frac{1}{N_s}\sum_{j=1}^{N_s}\delta(\boldsymbol{c}-\boldsymbol{c}^{(j)}) \tag{12-2}$$

式中：$\delta(\boldsymbol{c}-\boldsymbol{c}^{(j)})$为$\boldsymbol{c}^{(j)}$处的单位脉冲质量函数。

将式(12-2)代入状态预测方程(5-51)，得

$$\hat{\boldsymbol{x}}_{n|n-1} = \frac{1}{N_s}\sum_{j=1}^{N_s}\int\tilde{\boldsymbol{f}}(\boldsymbol{c})\delta(\boldsymbol{c}-\boldsymbol{c}^{(j)})\mathrm{d}\boldsymbol{c} = \frac{1}{N_s}\sum_{j=1}^{N_s}\tilde{\boldsymbol{f}}(\boldsymbol{c}^{(j)}) \tag{12-3}$$

当时 $N_s\to\infty$，由抽样均值得的逼近积分非常准确（见参考文献[1]）。注意，系数 $1/N_s$是每个样本的重要性权值，而不是概率。用归一化直方图表示指定区域，概率由该区域样本归一化密度得。

利用式(5-53)，式(12-3)变成

$$\boldsymbol{x}_{n|n-1} = \frac{1}{N_s}\sum_{j=1}^{N_s}\tilde{\boldsymbol{f}}(\hat{\boldsymbol{x}}_{n-1|n-1}+\boldsymbol{D}_{n-1|n-1}\boldsymbol{c}^{(j)}) \tag{12-4}$$

使用$\boldsymbol{\chi}_{n-1|n-1}$的标准定义，可有

$$\boldsymbol{\chi}_{n-1|n-1}^{(j)} = \hat{\boldsymbol{x}}_{n-1|n-1} + \boldsymbol{D}_{n-1|n-1}\boldsymbol{c}^{(j)} \tag{12-5}$$

定义 $\omega_j \triangleq 1/N_s, \forall j$ 将式(12－4)重写为

$$\hat{\boldsymbol{x}}_{n|n-1} = \sum_{j=1}^{N_s} w_j \tilde{\boldsymbol{f}}(\boldsymbol{\chi}_{n-1|n-1}^{(j)}) \tag{12-6}$$

其状态预测方程与所有 Sigma 点卡尔曼滤波器具有相同形式。

其余的蒙特卡洛卡尔曼滤波器(MCKF)预测方程由 Sigma 点预测方程给出，即

$$\boldsymbol{P}_{n|n-1}^{xx} = \sum_{j=0}^{N_s} w_j[\boldsymbol{f}(\boldsymbol{\chi}_{n-1|n-1}^{(j)}) - \hat{\boldsymbol{x}}_{n|n-1}] \cdot \tag{12-7}$$

$$[\boldsymbol{f}(\boldsymbol{\chi}_{n-1|n-1}^{(j)}) - \hat{\boldsymbol{x}}_{n|n-1}]^{\mathrm{T}} + \boldsymbol{Q} \tag{12-8}$$

$$\hat{z}_{n|n-1} = \sum_{j=0}^{N_s} w_j \boldsymbol{h}(\boldsymbol{\chi}_{n|n-1}^{(j)}) \tag{12-9}$$

$$\boldsymbol{P}_{n|n-1}^{zz} = \sum_{j=0}^{N_s} w_j[\boldsymbol{h}(\boldsymbol{\chi}_{n|n-1}^{(j)}) - \hat{\boldsymbol{x}}_{n|n-1}] \cdot \tag{12-10}$$

$$[\boldsymbol{h}(\boldsymbol{\chi}_{n|n-1}^{(j)}) - \hat{\boldsymbol{x}}_{n|n-1}]^{\mathrm{T}} + \boldsymbol{R} \tag{12-11}$$

$$\boldsymbol{P}_{n|n-1}^{xz} = \sum_{j=0}^{N_s} w_j[\boldsymbol{f}(\boldsymbol{\chi}_{n-1|n-1}^{(j)}) - \hat{\boldsymbol{x}}_{n|n-1}] \cdot \tag{12-12}$$

$$[\boldsymbol{h}(\boldsymbol{\chi}_{n|n-1}^{(j)}) - \hat{\boldsymbol{x}}_{n|n-1}]^{\mathrm{T}} \tag{12-13}$$

在后三个预测方程中，$\boldsymbol{\chi}_{n|n-1}^{(j)}$可写为

$$\boldsymbol{\chi}_{n|n-1}^{(j)} = \hat{\boldsymbol{x}}_{n|n-1} + \boldsymbol{D}_{n|n-1}\boldsymbol{c}^{(j)} \tag{12-14}$$

式中：$\boldsymbol{c}^{(j)}$ 为由$\mathcal{N}(\boldsymbol{c};\boldsymbol{0},\boldsymbol{I})$抽取的一组新的蒙特卡洛样本。所有这些预测方程实际上都是非线性函数的样本均值和协方差。

注意，MCKF 中从不重复使用蒙特卡洛样本。本书第Ⅲ部分将给出一些非高斯密度估计方法，反复重用蒙特卡洛样本，只是在每次迭代中更新它们的权重。这就是术语粒子滤波器的来历，每个蒙特卡洛样本都被当作是一个粒子，在迭代中传播。

12.1 蒙特卡洛卡尔曼滤波器

表 12－1 概括了 MCKF 过程。

表 12－1　蒙特卡洛卡尔曼滤波器过程(译者注:原文此处有误)

步骤	内容	公式
步骤(1)	滤波器初始化	设 $0 \leqslant \omega_0 \leqslant 1$ 初始化 $\hat{x}_0$ 和 $\boldsymbol{P}_0^{xx}$
步骤(2)	生成状态矢量样本预测	$w_j = 1/N_s \quad j = 1,\cdots,N_s$ $\boldsymbol{c}^{(j)} \sim \mathcal{N}(\boldsymbol{c};\boldsymbol{0},\boldsymbol{I})\ j = 1,\cdots,N_s$ $\boldsymbol{\chi}_{n-1\mid n-1}^{(j)} = \hat{\boldsymbol{x}}_{n-1\mid n-1} + \boldsymbol{D}_{n-1\mid n-1}\boldsymbol{c}^{(j)} \quad j = 1,\cdots,N_s$ $\hat{\boldsymbol{x}}_{n\mid n-1} = \sum_{j=1}^{N_s} w_j \boldsymbol{f}(\boldsymbol{\chi}_{n-1\mid n-1}^{(j)})$ $\boldsymbol{P}_{n\mid n-1}^{xx} = \sum_{j=0}^{N_s} w_j [\boldsymbol{f}(\boldsymbol{\chi}_{n-1\mid n-1}^{(j)}) - \hat{\boldsymbol{x}}_{n\mid n-1}] \cdot [\boldsymbol{f}(\boldsymbol{\chi}_{n-1\mid n-1}^{(j)}) -$ $\hat{\boldsymbol{x}}_{n\mid n-1}]^{\mathrm{T}} + \boldsymbol{Q}$
步骤(3)	生成观测值样本预测	$\boldsymbol{c}^{(j)} \sim \mathcal{N}(\boldsymbol{c};\boldsymbol{0},\boldsymbol{I}) \quad j = 1,\cdots,N_s$ $\boldsymbol{\chi}_{n\mid n-1}^{(j)} = \hat{\boldsymbol{x}}_{n\mid n-1} + \boldsymbol{D}_{n\mid n-1}\boldsymbol{c}^{(j)},\ j = 1,\cdots,N_s$ $\hat{z}_{n\mid n-1} = \sum_{j=0}^{N_s} w_j \boldsymbol{h}(\chi_{n\mid n-1}^{(j)})$ $\boldsymbol{P}_{n\mid n-1}^{zz} = \sum_{j=0}^{N_s} w_j [\boldsymbol{h}(\boldsymbol{\chi}_{n\mid n-1}^{(j)}) - \hat{z}_{n\mid n-1}][\boldsymbol{h}(\boldsymbol{\chi}_{n\mid n-1}^{(j)}) - \hat{z}_{n\mid n-1}]^{\mathrm{T}} + \boldsymbol{R}$ $\boldsymbol{P}_{n\mid n-1}^{xz} = \sum_{j=0}^{N_s} w_j [\boldsymbol{f}(\boldsymbol{\chi}_{n-1\mid n-1}^{(j)}) - \hat{\boldsymbol{x}}_{n\mid n-1}] \cdot$ $[\boldsymbol{h}(\boldsymbol{\chi}_{n\mid n-1}^{(j)}) - \hat{z}_{n\mid n-1}]^{\mathrm{T}}$
步骤(4)	卡尔曼滤波器更新	$\boldsymbol{K}_n \triangleq \boldsymbol{P}_{n\mid n-1}^{xz}(\boldsymbol{P}_{n\mid n-1}^{zz})^{-1}$ $\hat{\boldsymbol{x}}_{n\mid n} = \hat{\boldsymbol{x}}_{n\mid n-1} + \boldsymbol{K}_n(z_n^{o} - \hat{z}_{n\mid n-1})$ $\boldsymbol{P}_{n\mid n}^{xx} = \boldsymbol{P}_{n\mid n-1}^{xx} - \boldsymbol{K}_n \boldsymbol{P}_{n\mid n-1}^{zz} \boldsymbol{K}_n^{\mathrm{T}}$
步骤(5)	存储结果 返回步骤(2)	将 $\hat{\boldsymbol{x}}_{n\mid n}$ 和 $\boldsymbol{P}_{n\mid n}^{xx}$ 存储到跟踪文件中

参考文献

[1] Doucet A, deFreitas N, Gordon N. *Sequential Monte Carlo Methods in Practice.* Springer, 2001.

第13章
高斯卡尔曼滤波器总结

本章将总结第6~第12章中提出的高斯分布卡尔曼滤波器的重要特性。首先，分析LKF和EKF，给出每种算法的过程流图；然后，分析Sigma点类卡尔曼滤波器，给出通用的过程框图，并通过两个列表，定义每种滤波器对应的Sigma点和权值；最后，从处理效率的角度，对每种Sigma点卡尔曼滤波器进行性能特征的评价。

13.1 解析卡尔曼滤波器

在许多动态系统中，动态方程或者观测方程是线性的。本书中大多数案例也符合这一情况。

在某些情况下，动态方程是线性的，而观测方程是高度非线性的。这时，通常用式(6-3)和式(6-6)线性卡尔曼滤波器预测方程来代替其他的预测方法，进行状态预测和状态协方差预测。对于另外一些跟踪问题，当观测方程为线性时，动态方程可能为非线性。对于这类问题，处理方法与上面相反，即使用非线性方法进行状态预测，而采用式(6-9)~式(6-12)线性卡尔曼滤波器预测来进行观测值预测。最后，如果动态方程和观测方程都是线性的，则使用LKF预测方程。

图13-1给出了LKF过程框图。

当动态方程和/或观测方程为非线性时，必须使用第7~第12章中给出的非线性方法。过去50年中，使用最多的方法是EKF。在第7章中已经说明，EKF为非线性函数的一阶泰勒多项式展开得的卡尔曼滤波器，形式与LKF相同，但是需要额外计算雅可比矩阵，图13-2给出了此方法的框图。第7章还给出另外一种二阶扩展卡尔曼滤波器，但需要额外计算Hessian矩阵，这通常很难实现。

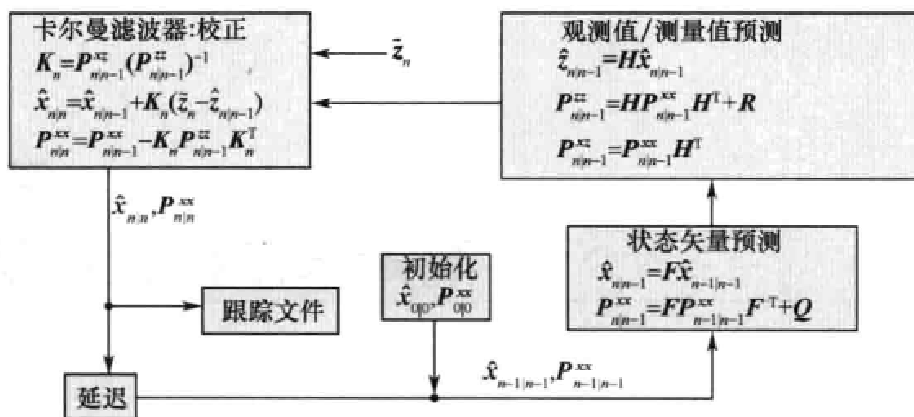


图 13-1　线性卡尔曼滤波器过程框图

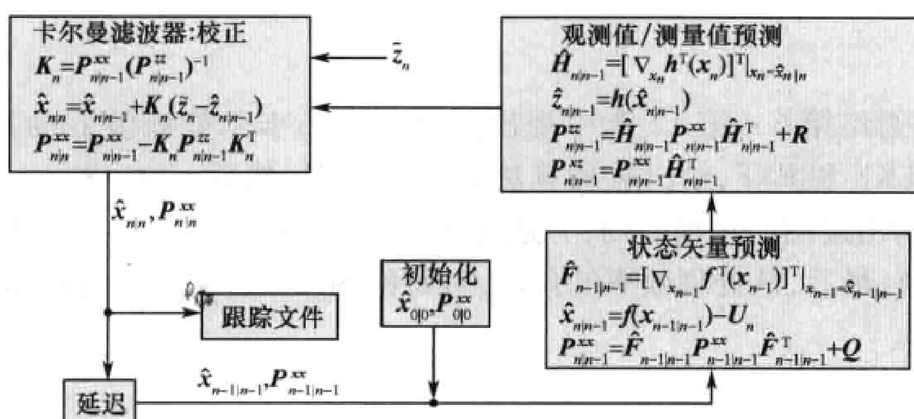


图 13-2　EKF 过程框图

13.2　Sigma 点卡尔曼滤波器

第 8 章中提出用中心有限差分近似取代雅可比矩阵和 Hessian 矩阵的计算，以得不同形式的 FDKF 滤波器。由式(8-41)、式(8-43)和式(8-47)给出了标准 FDKF 的协方差预测方程，该方程与其他的 Sigma 点方法预测方程在形式上不同。而改进后的 FDKF 滤波器的协方差预测方程式(8-49)、式(8-50)和式(8-51)，与 UKF 的预测方程是完全相同的。因此，在总结 Sigma 点卡尔曼滤波器时，没有将 FDKF 包含在内。

Sigma 点卡尔曼滤波器——UKF、SSKF、GHKF 和 MCKF，都具有相似的结构，因此可以用图 13-3 中的通用流程进行处理，只是 $c^{(j)}$ 和 ω_j 取值有所不同。表 13-1 和表 13-2 分别列出了每种滤波器中 $c^{(j)}$ 和 ω_j 的取值。基于图 13-3 的过程框图，编写一个通用的 Sigma 点卡尔曼滤波器子程序相对比较容易，该程序可调用二级子程序，从表 13-1 和表 13-2 中选取合适的 Sigma 点和权值。

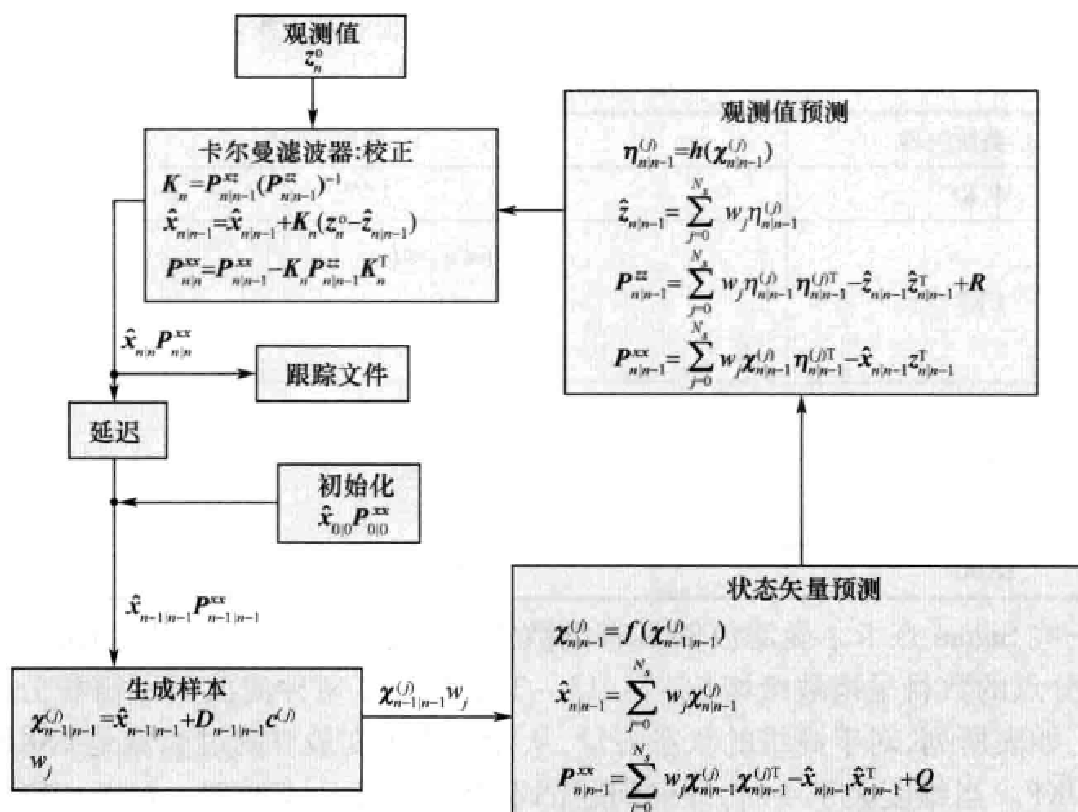


图 13-3 通用 Sigma 点卡尔曼滤波器过程框图

表 13-1 Sigma 点卡尔曼滤波器数据汇总:第 1 部分
—用于 Sigma 点卡尔曼滤波器的 c 的形式

卡尔曼滤波器	$c^{(j)}$ 的形式
MCKF	$\boldsymbol{c}^{(j)} \sim \mathcal{N}(\boldsymbol{c};\boldsymbol{0},\boldsymbol{I}) \quad j=1,\cdots,N_s$
UKF	$\boldsymbol{c}^{(j)} = \sqrt{\dfrac{n_x}{1-w_0}}\boldsymbol{r}^{(j)} \quad j=0,1,\cdots,2n_x$ $\boldsymbol{r}^{(j)} = \begin{matrix} [\boldsymbol{0}] \in \mathbf{R}^{n_x} & j=0 \\ [1] \in \mathbf{R}^{n_x} & j=1,2,\cdots,2n_x \end{matrix}$
SSKF	$\boldsymbol{c}_{\mathbf{R}^{n_x}}^{(j)} = \begin{cases} [\boldsymbol{0}] \in \mathbf{R}^{n_x} & j=0 \\ \begin{bmatrix} \boldsymbol{c}_{\mathbf{R}^{n_x-1}}^{(j)} \\ -\dfrac{1}{\sqrt{n_x(n_x+1)w_1}} \end{bmatrix} & j=1,2,\cdots,n_x \\ \begin{bmatrix} \boldsymbol{0} \in \mathbf{R}^{n_x-1} \\ \dfrac{n_x}{\sqrt{n_x(n_x+1)w_1}} \end{bmatrix} & j=n_x+1 \end{cases}$
GHKF	$\boldsymbol{c}^{(j)} = \sqrt{3}\boldsymbol{r}^{(j)} \quad \boldsymbol{r}^{(j)}$ 见式(11-81)

表 13-2 Sigma 点卡尔曼滤波器数据汇总:第 2 部分
—Sigma 点权值的形式

卡尔曼滤波器	w_j 的形式
MCKF	$1/N_s$
UKF	$\begin{cases} 0 \leqslant w_0 \leqslant 1 \\ w_j = \dfrac{1-w_0}{2n_x} \quad j=1,\cdots,2n_x \end{cases}$
SSKF	$\begin{cases} 0 \leqslant w_0 \leqslant 1 \\ w_j = \dfrac{1-w_0}{n_x+1} \quad j=1,\cdots,n_x+1 \end{cases}$
GHKF	w_j 见式(11-82)

一些 Sigma 点卡尔曼滤波器具有维数灾难问题,即当状态矢量的维数增加时,积分点的数目呈指数级增长。表 13-3 中列举了每种滤波器所需积分点的数目。如表所列,对于高维的状态矢量,从计算的角度最好的过滤器是 SSKF,其次是 UKF。当维度大于 4 时,任何问题都不会用 GHKF 来处理。对于高度非线性动态方程,每个滤波器的精度通常是由积分多项式展开的阶数所决定。因此,MCKF 的精度最高,只要随机样本数量足够高,非线性函数中不会用到任何近似;其次是 GHKF,其使用最高的多项式阶数;第三是 UKF,其使用三阶多项式;最后,最不精确的是 SSKF。因此,对于给定的估计问题,最好对所有滤波器进行试验,然后在计算量约束内,选择均方根误差最小的滤波器。Wu 等[1]的论文中可以找到关于上述滤波器以及相关其他滤波器在精度、效率和稳定性方面的讨论。

表 13-3 Sigma 点卡尔曼滤波器所需积分点数目的比较

	采样点个数	状态矢量维数				
		1	2	4	8	16
UKF	$2n_x+1$	3	5	9	17	33
SSKF	n_x+2	3	4	6	10	18
GHKF	3^{n_x}	3	9	81	6561	43046721
MCKF	N_s	N_s	N_s	N_s	N_s	N_s

列表 13-1、表 13-2 给出任意状态矢量维度下,为所有的 Sigma 点卡尔曼滤波器生成集合$\{w_i, c^{(i)}, i\cdots 1,\cdots,N_s\}$的 MATLAB 子程序。第一个子程序分别根据表 13-1 和表 13-2 中的公式,计算权值和 Sigma 点。该程序有 4 个输入,分别为:标志位(t_flag)指明是针对哪种 Sigma 点卡尔曼滤波器计算 Sigma 点和权值;状态矢量维数($L=n_x$)为 Sigma 点(或采样点))的个数;LL,由表 13-3 计算得;最后是数值流,为随机数值发生器的流。第二个子程序实现了高斯-厄米特矢量点的生成,由美国约翰·霍普金斯大学应用物理实验室 Neill Bryant Bel-

tran 编写,函数输出长度为 u、维数为 n 的 Sigma 点。为实现这一功能,用到了二进制数及其组合。首先,生成所有可能二进制数,其位数与维度相同;其次,对这些二进制数字进行筛选,仅保留有 u 个“1”的二进制数;再次,生成一个负值的掩码,其中包括 1 和 -1 所有可能组合;最后,结合“形状掩码”和“负值掩码”,找出高斯-厄米特矢量点的所有可能组合。

列表 13-1　生成权值和 Sigma 点

```
function [w,c] = WeightsSigmaPoints(t_flag,L,LL,Stream)
w = zeros(1,LL);
switch tflag
    case{1}         % UKF
        w(1) = 0.25;
        w(2:LL) = (1 - w(1))/(2 * L);
        r = zeros(2 * L + 1,L);
        r(2:L + 1,:) = eye(L);
        r(L + 2:2 * L + 1,:) = - eye(L);
        c = sqrt(L/(1 - w(1))) * r;
    case{2}         % SSKF
        w(1) = 0.25;

        w(2:LL) = (1 - w(1))/(L + 1);
        c = zeros(LL,L);
        c(2,1) = 1/sqrt(2 * w(2));
        c(3,1) = -1/sqrt(2 * w(2));
        for n = 2:L
            for q = 1:n
                c(q + 1,1:n) = [c(q + 1,1:n - 1)1/…
                    sqrt(n * (n + 1) * w(2))];
            end
                c(n + 2,1:n) = [zeros(1,n - 1) - n/…
                    sqrt(n * (n + 1) * w(2))];
        end
    case{3}         % GHKF
        fact1 = 2/3;
        fact2 = 1/6;
        r = zeros(LL,L);
        w(1) = fact1^L;
        q = 1;
        for u = 1:L
            rr = vectorPointGenerator(u,L);
```

```
            k = size(rr,1);
            r(q +1:q + k,:) = rr;
            w(q +1:q + k) = (fact1^(L - u)) * (fact2^u);
            q = q + k;
        end
        c = sqrt(3) * r;
    case{4}        % MCKF
        w = ones(1,LL) / LL;
        c = randn(Stream,LL,L);
end
```

列表 13 - 2 矢量点发生器

```
function out = vectorPointGenerator(u,n)
% Written by
% Neill Bryant Beltran
% 8 November 2010
i f u = = 0
    out = zeros(1,n);
    return
end
% Find a11 binary numbers with same number
% of digits as there are dimensions,n
shape_mask = dec2bin(2^n - 1: - 1:1) = = '1';
shape_size = size(shape_mask);
% Remove a11 vectors that do not have u '1's
for ii = shape_size(1): - 1:1
    i f nnz(shape_mask(ii,:)) ~ = u
        shape_mask(ii,:) = [];
    end
end
shape_size = size(shape_mask);
% Create the negative mask
neg_mask = (dec2bin(0:2^u - 1) = = '1') * 1.0;
neg_mask(neg_mask = = 0) = - 1;
neg_mask = - 1 * neg_mask;
neg_size = size(neg_mask);
out = zeros(shape_size(1) * neg_size(1),n);
% Combine the shape mask and negative mask
for ii =1:neg_size(1)
    for jj =1:shape_size(1)
```

```
        kk =1;
        for ll =1:n
            if ~shape_mask(jj,ll)
                continue
            end
            out((ii -1) * shape_size(1) + jj,ll) = ...
                shape_mask(jj,ll) * neg_mask(ii,kk);
            kk = kk + 1;
        end
    end
end
```

在文献[2－7]中可以找到计算高斯加权积分的其他算法。文献中讨论的许多算法的实际应用类似于前面提出的 Sigma 点卡尔曼滤波器。与第 8～第 12 章给出的方法相比，它们大部分阶数比较高，因此与 UKF 相比需要更多积分点，所需点的数目介于 UKF 和 GHKF 之间。在 Candy[8]的书中，给出了更为经典的 Sigma 点卡尔曼滤波器的推导，采用回归方法，类似于卡尔曼滤波器的原始推导。

13.3 一种更实用的运用卡尔曼滤波器族的方法

通过各种案例分析，对式(3－17)和式(3－18)给出的动态模型和观测模型进行仔细检查，发现存在一个层次模型。表 13－4 给出了这个层次模型。

表 13－4　动态模型和观测模型层次

组合	动态模型	观测模型
线性/线性	$\boldsymbol{x}_n = \boldsymbol{F}\boldsymbol{x}_{n-1} + \boldsymbol{v}_{n-1}$	$z_n = \boldsymbol{H}\boldsymbol{x}_n + \boldsymbol{w}_n$
线性/非线性	$\boldsymbol{x}_n = \boldsymbol{F}\boldsymbol{x}_{n-1} + \boldsymbol{v}_{n-1}$	$z_n = \boldsymbol{f}(\boldsymbol{x}_n) + \boldsymbol{w}_n$
非线性/线性	$\boldsymbol{x}_n = \boldsymbol{f}(\boldsymbol{x}_{n-1}) + \boldsymbol{v}_{n-1}$	$z_n = \boldsymbol{H}\boldsymbol{x}_n + \boldsymbol{w}_n$
非线性/非线性	$\boldsymbol{x}_n = \boldsymbol{f}(\boldsymbol{x}_{n-1}) + \boldsymbol{v}_{n-1}$	$z_n = \boldsymbol{f}(\boldsymbol{x}_n) + \boldsymbol{w}_n$

基于此模型的层次结构，可以认为动态预测过程是完全独立于观测预测过程的。

这使人们能够选用每种模型最适合的卡尔曼滤波方法。例如，对于线性/非线性情况，可以使用线性卡尔曼滤波器方程进行动态预测，而采用非线性卡尔曼滤波器方法进行观测预测。同样，对于非线性/非线性情况，当状态矢量 $\boldsymbol{x}_n$ 为高维而观测矢量 z_n 是低维时，可以使用 UKF 进行动态预测(因为 GHKF 不实用)，而使用 GHKF 进行观测预测(因为它通常是最准确的)。上述讨论表明了卡尔曼滤波器架构的灵活性，在解决具体问题时，可以建立最高效的估计过程，进行

实际应用。当然，从软件实现的角度来看，这也许并不是最高效的方法。

参考文献

[1] Wu Y, Hu D, Wu M, Hu X. A numerical – integration perspective on Gaussian filters. *IEEE Trans. Sig. Proc*, 2006, 54(8): 2910 – 2921.

[2] McNamee J, Stenger F. Construction of fully symmetric numerical integration formulas. *Numerische Math*, 1967, 10: 327 – 344.

[3] Haber S. Numerical evaluation of multiple integrals. *SIAM Rev*, 1970, 12(4): 481 – 526.

[4] Dunavant D A. Efficient symmetrical cubature rules for complete polynomials of higher degree over the unit cube. *Int. J. Numer. Methods Eng*, 1986, 23: 397 – 407.

[5] Genz A, Monahan J. A stochastic algorithm for higher – dimensional integrals over unbounded regions with Gaussian weights. *J. Comput. Appl. Math*, 1999, 112: 71 – 81.

[6] Cools R. Advances in multidimensional integration. *J. Comput. Appl. Math*, 2002, 149: 1 – 12.

[7] Lu J, Darmofal D L. Higher – dimensional integration with Gaussian weight for applications in probabilistic design. *SIAM J. Sci. Comput*, 2004, 26(2): 613 – 624.

[8] Candy JV. Bayesian Signal Processing: Classical, Modern, and Particle Filtering Methods. Hoboken NJ: Wiley, 2009.

第14章 卡尔曼滤波器族的性能度量

本章将讨论一些通用的方法来度量第6～第12章提出的卡尔曼滤波器的性能。性能度量方法分为两类,第一类适用于无法得到实际轨迹数据或者针对特殊场景,例如在对少量的现场试验的分析中,可得到一些实际轨迹数据,但统计意义并不明显。对于这种情况,只能将滤波器估计轨迹协方差矩阵作为输入,生成连续的跟踪误差椭圆,然后进行性能度量。第二类适用于用仿真来生成一个完全已知的目标轨迹,这样就可以进行较多的性能度量。这是因为可以生成大量的蒙特卡洛观测数据集,这样就能实施各种跟踪性能度量方法,如均方根(RMS)误差、Cramer－Rao下界(CRLB)和发散轨迹比率等。本章将对所有的性能度量方法以及每种方法相关的具体实施问题进行讨论。

14.1 误差椭圆

跟踪估计滤波器的应际应用中,当真实状态为未知,比如在一个现场系统中跟踪一艘真正的船舶,或者当只进行了少量的现场试验、仅知道大约的真实状态时,最好的衡量跟踪滤波器性能的方式是利用滤波器的估计协方差矩阵输出来生成并绘制滤波跟踪状态的误差椭圆。由于屏幕显示所固有的二维特性,为清楚描述,这里仅采用二维笛卡儿坐标系统下的位置。

在任意离散时刻t_n,跟踪状态$\boldsymbol{x}_n$的概率密度函数为高斯分布,即

$$\boldsymbol{x}_n \sim \mathcal{N}(\boldsymbol{x}_n;\hat{\boldsymbol{x}}_{n|n},\boldsymbol{P}_{n|n}^{xx}) \quad (14-1)$$

式中:$(\hat{\boldsymbol{x}}_{n|n},\boldsymbol{P}_{n|n}^{xx})$为笛卡儿坐标系下的卡尔曼滤波器更新方程。现在,不考虑时间下标,只检查由$\boldsymbol{x}=[r^x,r^y]^{\mathrm{T}}$指定的二维位置分量。如图2－9所示,常数概率(仿射变换之前)的椭圆轮廓一般沿主轴方向拉长,也可能是倾斜的,这样椭圆的轴与状态位置矢量的轴不平行。下面的步骤中将回答问题,即给定一个包含

概率，如何在与状态矢量相同的坐标系中绘制误差椭圆。这里包含概率相当于常数概率的轮廓。

14.1.1 正则椭圆

位置分量概率分布的指数可以表示为

$$s = (\boldsymbol{x} - \hat{\boldsymbol{x}})^{\mathrm{T}} \boldsymbol{P}^{-1} (\boldsymbol{x} - \hat{\boldsymbol{x}}) \tag{14-2}$$

式中：$\boldsymbol{P}$ 为位置分量的误差协方差矩阵。误差椭圆定义为

$$(\boldsymbol{x} - \hat{\boldsymbol{x}})^{\mathrm{T}} \boldsymbol{P}^{-1} (\boldsymbol{x} - \hat{\boldsymbol{x}}) \leqslant C^2 \tag{14-3}$$

式中：C^2 为包含区域，取决于要得到的包含概率 P_c，后面将给出 P_c 的定义。

首先，将式(14-3)改写为椭圆方程的标准形式。对 $\boldsymbol{P}$ 进行特征值分解，得

$$\boldsymbol{P} = \boldsymbol{V}\boldsymbol{D}\boldsymbol{V}^{\mathrm{T}} \tag{14-4}$$

式中：$\boldsymbol{D}$ 为一个包含 $\boldsymbol{P}$ 的特征值的 2×2 对角矩阵，即

$$\boldsymbol{D} \triangleq \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \tag{14-5}$$

而 $\boldsymbol{V}$ 是一个由 $\boldsymbol{P}$ 的两个特征矢量构成的 2×2 矩阵，即

$$\boldsymbol{V} \triangleq [\boldsymbol{v}_1, \boldsymbol{v}_2] \tag{14-6}$$

式中：$\boldsymbol{v}_1$ 和 $\boldsymbol{v}_2$ 为 2×1 列矢量。很容易证明，对于一个对称、非奇异矩阵 $\boldsymbol{P}$，有

$$\boldsymbol{P}^{-1} = \boldsymbol{V}\boldsymbol{D}^{-1}\boldsymbol{V}^{\mathrm{T}}$$

式(14-3)可重写为

$$(\boldsymbol{x} - \hat{\boldsymbol{x}})^{\mathrm{T}} \boldsymbol{V}\boldsymbol{D}^{-1}\boldsymbol{V}^{\mathrm{T}} (\boldsymbol{x} - \hat{\boldsymbol{x}}) \leqslant \boldsymbol{C}^2 \tag{14-7}$$

令

$$\boldsymbol{y} \triangleq \boldsymbol{V}^{\mathrm{T}} (\boldsymbol{x} - \hat{\boldsymbol{x}}) \tag{14-8}$$

将式(14-7)重写为

$$\boldsymbol{y}^T \boldsymbol{D}^{-1} \boldsymbol{y} = \frac{y_1^2}{\lambda_1} + \frac{y_2^2}{\lambda_2} \leqslant C^2 \tag{14-9}$$

选取 λ_1 作为最大的特征值，并做如下定义，即

$$a \triangleq C\sqrt{\lambda_1} \tag{14-10}$$

$$b \triangleq C\sqrt{\lambda_2} \tag{14-11}$$

选取式(14-7)中的等式，得正则椭圆方程为

$$\frac{y_1^2}{a^2} + \frac{y_2^2}{b^2} = 1 \tag{14-12}$$

14.1.2 确定 $\boldsymbol{P}$ 的特征值

为确定对称矩阵 $\boldsymbol{P}$ 的特征值与特征矢量,检验特征值方程

$$\boldsymbol{P}\boldsymbol{v} = \lambda \boldsymbol{v} \tag{14-13}$$

$$[\boldsymbol{P} - \lambda \boldsymbol{I}]\boldsymbol{v} = \boldsymbol{0} \tag{14-14}$$

多项式 $f(\lambda) = |\boldsymbol{P} - \lambda\boldsymbol{I}|$ 被称为 $\boldsymbol{P}$ 的特征多项式,$\bar{f}(\lambda)=0$ 的根即为特征值 λ_i,有

$$|\boldsymbol{P} - \lambda\boldsymbol{I}| = \begin{vmatrix} P_{11}-\lambda & P_{12} \\ P_{12} & P_{22}-\lambda \end{vmatrix} = 0 \tag{14-15}$$

得如下解

$$\lambda_1 = \frac{P_{11}+P_{22}}{2} + \frac{1}{2}\sqrt{(P_{11}-P_{22})^2 + 4P_{12}^2} \tag{14-16}$$

$$\lambda_2 = \frac{P_{11}+P_{22}}{2} - \frac{1}{2}\sqrt{(P_{11}-P_{22})^2 + 4P_{12}^2} \tag{14-17}$$

14.1.3 确定误差椭圆旋转角度

对于对应特征值 λ_1 的特征矢量 $\boldsymbol{v}_1$,式(14-14)变为

$$\begin{bmatrix} P_{11}-\lambda_1 & P_{12} \\ P_{12} & P_{22}-\lambda_1 \end{bmatrix}\boldsymbol{v}_1 = \boldsymbol{0} \tag{14-18}$$

令

$$\boldsymbol{v}_1 \triangleq \begin{bmatrix} \xi_{11} \\ \xi_{12} \end{bmatrix} \tag{14-19}$$

$$\boldsymbol{v}_2 \triangleq \begin{bmatrix} \xi_{21} \\ \xi_{22} \end{bmatrix} \tag{14-20}$$

式中:(ξ_{11},ξ_{12})和(ξ_{21},ξ_{22})分别为 $\boldsymbol{v}_1$ 和 $\boldsymbol{v}_2$ 的笛卡儿坐标分量。已知 ξ_{11},通过式(14-18)求 ξ_{12},得

$$\xi_{12} = -\frac{P_{12}}{P_{22}-\lambda_1}\xi_{11} \tag{14-21}$$

图 14-1 给出了椭圆相对于笛卡儿轴的旋转角。由图可知,误差椭圆旋转角度为

$$\varphi = \tan^{-1}\left(\frac{\xi_{12}}{\xi_{11}}\right) = \tan^{-1}\left(-\frac{P_{12}}{\lambda_1 - P_{22}}\right) \tag{14-22}$$

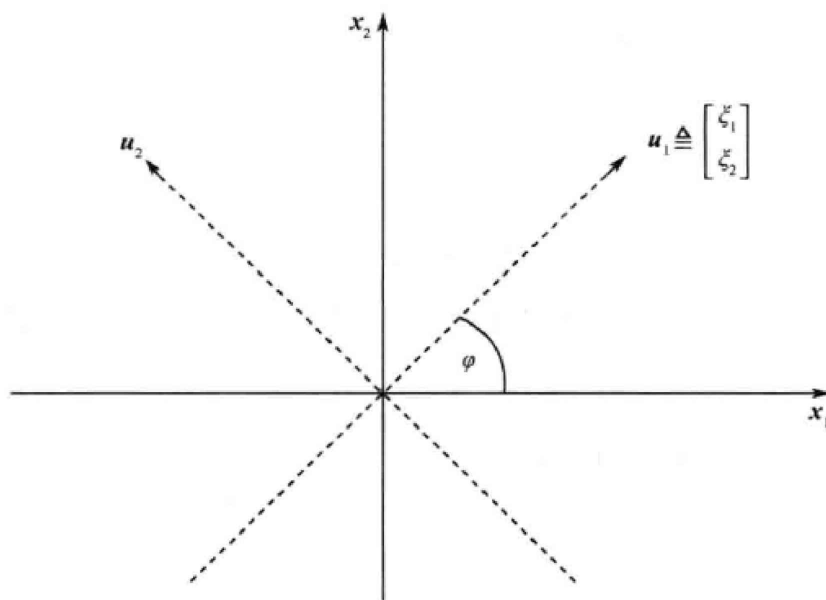


图 14－1　误差椭圆旋转角度

14.1.4　确定包含区域

由于 $\boldsymbol{x}$ 的分量是联合高斯分布的，如式（14－1）可知，$\boldsymbol{y}$ 的分量同样如此，因此有

$$\boldsymbol{y} \sim \mathcal{N}(\boldsymbol{y};\boldsymbol{0},\boldsymbol{D}) \tag{14-23}$$

$\boldsymbol{y}$ 的概率密度函数可写为

$$p(\boldsymbol{y}) = \frac{1}{2\pi\sqrt{\lambda_1\lambda_2}} \mathrm{e}^{-\left[-\frac{y_1^2}{\lambda_1}+\frac{y_2^2}{\lambda_2}\right]} \tag{14-24}$$

因此，包含概率可以表示为

$$P_C = P\left\{\frac{y_1^2}{\lambda_1} + \frac{y_2^2}{\lambda_2} \leqslant C^2\right\} = \iint\limits_{\frac{y_1^2}{\lambda_1}+\frac{y_2^2}{\lambda_2} \leqslant C^2} p(\boldsymbol{y})\mathrm{d}\boldsymbol{y} \tag{14-25}$$

令

$$r^2 \triangleq \frac{y_1^2}{\lambda_1} + \frac{y_2^2}{\lambda_2} \leqslant C^2 \tag{14-26}$$

得

$$y_1 = \sqrt{\lambda_1} r\cos\varphi \tag{14-27}$$

$$y_2 = \sqrt{\lambda_2} r\sin\varphi \tag{14-28}$$

$$\mathrm{d}y_1\mathrm{d}y_2 = \sqrt{\lambda_1\lambda_2} r\mathrm{d}r\mathrm{d}\varphi \tag{14-29}$$

包含概率式（14－25）变为

$$PC = \frac{1}{2\pi}\int_0^{2\pi} \mathrm{d}\varphi \int_0^C r\mathrm{e}^{-\frac{1}{2}r^2}\mathrm{d}r = \int_0^C r\mathrm{e}^{-\frac{1}{2}r^2}\mathrm{d}r = 1 - \mathrm{e}^{-\frac{1}{2}C^2} \tag{14-30}$$

对 C 进行求解,得

$$C = \sqrt{-2\ln(1 - PC)} \tag{14-31}$$

因此,选定一个包含概率就完全决定了 C 的取值。

14.1.5 误差椭圆参数化绘图

一个正则椭圆(形式为式(14-12))可以表示为点$(x_e(\boldsymbol{\theta}), y_e(\boldsymbol{\theta}))$的轨迹,即

$$x_e(\boldsymbol{\theta}) = x_c + a\cos\boldsymbol{\theta}\cos\varphi - b\sin\boldsymbol{\theta}\sin\varphi \tag{14-32}$$

$$y_e(\boldsymbol{\theta}) = y_c + a\cos\boldsymbol{\theta}\sin\varphi + b\sin\boldsymbol{\theta}\cos\varphi \tag{14-33}$$

这里(x_c, y_c)为椭圆中心,φ 为笛卡儿坐标系 x 轴与椭圆主轴的夹角,参数矢量 $\boldsymbol{\theta}$ 定义为行矢量 $\boldsymbol{\theta} \triangleq [0, 0.01, \cdots, 2\pi]$。

定义参数矩阵

$$\boldsymbol{Q}(\boldsymbol{\theta}) = \begin{bmatrix} \boldsymbol{q}_1(\boldsymbol{\theta}) \\ \boldsymbol{q}_2(\boldsymbol{\theta}) \end{bmatrix} = \begin{bmatrix} \cos\boldsymbol{\theta} \\ \sin\boldsymbol{\theta} \end{bmatrix} \tag{14-34}$$

由式(14-10)和式(14-11),可以证明

$$\begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix} = C\begin{bmatrix} \sqrt{\lambda_1} & 0 \\ 0 & \sqrt{\lambda_2} \end{bmatrix} = C\boldsymbol{D}^{1/2} \tag{14-35}$$

$$\boldsymbol{V} = \begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix} = \begin{bmatrix} \xi_{11} & \xi_{21} \\ \xi_{12} & \xi_{22} \end{bmatrix} \tag{14-36}$$

可以将椭圆轨迹的参数方程写成矢量-矩阵形式

$$x_e(\boldsymbol{\theta}) = \begin{bmatrix} x_e(\boldsymbol{\theta}) \\ y_e(\boldsymbol{\theta}) \end{bmatrix} = \begin{bmatrix} \hat{x} \\ \hat{y} \end{bmatrix} + \begin{bmatrix} \xi_{11}C\sqrt{\lambda_1}\boldsymbol{q}_1(\boldsymbol{\theta}) + \xi_{12}C\sqrt{\lambda_2}\boldsymbol{q}_2(\boldsymbol{\theta}) \\ \xi_{21}C\sqrt{\lambda_1}\boldsymbol{q}_1(\boldsymbol{\theta}) + \xi_{22}C\sqrt{\lambda_2}\boldsymbol{q}_2(\boldsymbol{\theta}) \end{bmatrix} \tag{14-37}$$

这里将椭圆中心设为$(\hat{x}, \hat{y})$。或者,可以将参数方程表示为

$$\boldsymbol{x}_e(\boldsymbol{\theta}) = \hat{\boldsymbol{x}} + \boldsymbol{V}C\boldsymbol{D}^{1/2}\boldsymbol{Q}(\boldsymbol{\theta}) \tag{14-38}$$

$$\hat{\boldsymbol{x}} \triangleq [\hat{x}, \hat{y}]$$

附注 14.1 可调用 MATLAB 函数$[V, D] = \mathrm{eig}(P)$分别得特征值 V 和特征矢量矩阵 $\boldsymbol{D}$。

在笛卡儿跟踪估计图基础上绘制生成误差椭圆数据的步骤如下：

（1）选取包含概率 Pc 的值，通常为 0.9，用式(14－31)计算 C 值；

（2）生成参数矢量 $\boldsymbol{\theta} \triangleq [0,0.01,\cdots,2\pi]$，除 0.01 之外，间隔也可以取其他任意值；

（3）由式(14－34)生成 $\boldsymbol{Q}(\boldsymbol{\theta})$；

（4）生成协方差矩阵 $\boldsymbol{P}^{xx}_{n|n}$ 的特征值 $\boldsymbol{V}$ 和特征矢量矩阵 $\boldsymbol{D}$；

（5）生成 $\boldsymbol{x}_e(\boldsymbol{\theta}) = \hat{\boldsymbol{x}}_{n|n} + \boldsymbol{V}C\boldsymbol{D}^{1/2}\boldsymbol{Q}(\boldsymbol{\theta})$。

14.1.6 误差椭圆举例

图 14－2 绘制了 DIFAR 实例中 UKF 跟踪输出中每 100 个数据点对应的误差椭圆。当真实航迹未知时，通过这种标绘方法，可以沿轨迹上各点进行跟踪估计，并评价估计的不确定性。

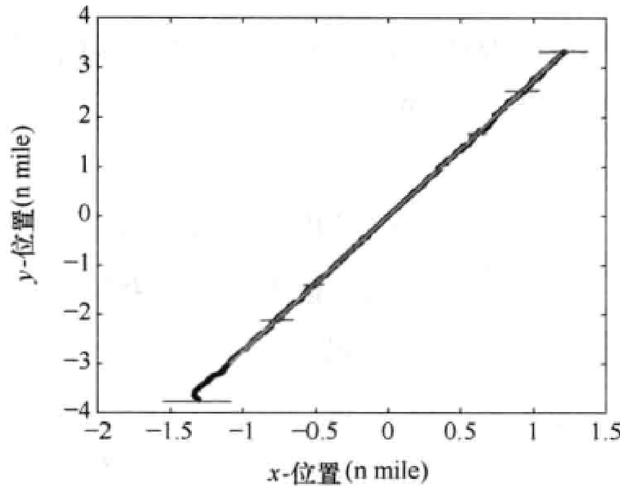


图 14－2　在 DIFAR 实例中 UKF 跟踪输出上叠加绘制的误差椭圆示例

14.2 均方根误差

利用式(3－16)，在 t_n 时刻，可以将状态协方差矩阵定义为

$$\boldsymbol{P}^{xx}_{n|n} = \varepsilon\{(\boldsymbol{x}_n - \hat{\boldsymbol{x}}_{n|n})(\boldsymbol{x}_n - \hat{\boldsymbol{x}}_{n|n})^{\mathrm{T}}\} \tag{14-39}$$

假设在仿真当中，$\boldsymbol{x}_n$ 已知，而 $\hat{\boldsymbol{x}}_{n|n}$ 是无偏的，对角元素 $\boldsymbol{P}^{xx}_{n|n}$ 的平方根可以由条件均方根误差（或者误差平方均值的平方根）来近似。

在第 4 章中曾强调，仿真的首要任务是对所有关注的时刻 $\{t_n, n = 0,1,\cdots,N\}$，生成一个真实的状态轨迹 $\{\boldsymbol{x}_n, n = 0,1,\cdots,N\}$。由真实状态轨迹可生成真实观测轨迹，在每个观测值中加入独立噪声，以得一组仿真观测值 $\{z_n, n = 0,1,\cdots,N\}$。为计算均方根误差，用蒙特卡洛方式生成多组观测值，这样就有了 M 组独立的观测值 $\{z_{n,m}; n = 0,1,\cdots,N; m = 1,2,\cdots,M\}$。注意这里的观测值噪声必须

独立于时间采样点（$z_{i,m}$必须独立于$z_{j,m}$，$\forall i \neq j$）和蒙特卡洛运行（$z_{n,l}$必须独立于$z_{n,k}$，$\forall l \neq k$）。

用给定的跟踪滤波器处理各组蒙特卡洛观测值，以产生一组跟踪文件输出$\hat{\boldsymbol{x}}_{n|n}(m)$和在$n$处的均方误差矩阵，得

$$\boldsymbol{S}(\hat{\boldsymbol{x}}_{n|n}) \triangleq \frac{1}{M}\sum_{m=1}^{M}[\boldsymbol{x}_n - \hat{\boldsymbol{x}}_{n|n}(m)][\boldsymbol{x}_n - \hat{\boldsymbol{x}}_{n|n}(m)]^{\mathrm{T}} \tag{14-40}$$

均方根误差可写为

$$\boldsymbol{e}_{\mathrm{RMS}}(\boldsymbol{x}_n) = \mathrm{diag}\left\{\left[\frac{1}{M}\sum_{m=1}^{M}[\boldsymbol{x}_n - \hat{\boldsymbol{x}}_{n|n}(m)][\boldsymbol{x}_n - \hat{\boldsymbol{x}}_{n|n}(m)]^{\mathrm{T}}\right]^{1/2}\right\} \tag{14-41}$$

这里 diag{·}表示“取对角元素并将其放入一个列矢量。”或者，换句话说，对于一个 4 维状态矢量$\boldsymbol{x}_n = [x_{1,n}, x_{2,n}, x_{3,n}, x_{4,n}]^{\mathrm{T}}$，有

$$\boldsymbol{e}_{\mathrm{RMS}}(\boldsymbol{x}_n) = \begin{bmatrix} e_{\mathrm{RMS}}(x_{1,n}) \\ e_{\mathrm{RMS}}(x_{2,n}) \\ e_{\mathrm{RMS}}(x_{3,n}) \\ e_{\mathrm{RMS}}(x_{4,n}) \end{bmatrix} \tag{14-42}$$

$$e_{\mathrm{RMS}}(x_{i,n}) \triangleq \left[\frac{1}{M}\sum_{m=1}^{M}[x_{i,n} - \hat{x}_{i,n|n}(m)]^2\right]^{1/2} \tag{14-43}$$

14.3 发散轨迹

在计算均方根误差性能时必须注意，因为即使是 100 次蒙特卡洛运行中的 1 个发散轨迹，也会造成均方根误差统计的偏差。可以将发散轨迹看作异常值，它不符合高斯假设，即均方根误差是标准偏差的无偏采样数据估计。不过，发散轨迹的判断是一个主观的决定，通常与实际场景相关。

对于 DIFAR 案例性能分析，如果x轴或y轴的估计值在最后时刻偏离真实值大于 2nmi，就判定估计轨迹为发散轨迹。由第 7～第 12 章内容可知，在 DIFAR 实例中，当信噪比低于 5dB 时，随着信噪比的降低，发散轨迹的数目增加。

14.4 Cramer－Rao 下界

有一个问题，给定场景（一组动态模型和观测模型以及状态轨迹）下，可得到的最好的（最小的）均方根误差是多少？因为总是存在观测噪声，当使用的动态模型中没有任何动态噪声时，就可以得到最好的均方根误差。

那么下界有什么用呢？评估可行的估计算法(跟踪器)的性能可能很困难，而下界会给出估计性能指标。在许多情况下，需要用性能下界来确定系统性能需求是不是能够实现。

本节大部分的内容是基于文献[1-3]及其所引用的参考文献。本节主要关注 Cramer-Rao 下界，但文献[2]超出这一界限，研究用各种不同的边界来解决问题，而不局限于本书的贝叶斯估计方法。

14.4.1 一维情况

考虑一维情况，如果过程(动态)噪声为零，则 x 是确定的。这样，唯一会导致 $\hat{x}$ 的误差的是观测噪声，可以用似然函数来表示。对于规范化的似然函数，有

$$\int p(z \mid x)\mathrm{d}z = 1 \tag{14-44}$$

$\hat{x}(z)$ 的期望值由下式给出，即

$$\int \hat{x}(z)p(z \mid x)\mathrm{d}z = x \tag{14-45}$$

以此为基础，由于 x 值确定，式(14-45)可以写成

$$\int (\hat{x}(z) - x)p(z \mid x)\mathrm{d}z = 0 \tag{14-46}$$

关于 x 求导，可推导出

$$\frac{\partial}{\partial x}\int (\hat{x} - x)p(z \mid x)\mathrm{d}z = 0 \tag{14-47}$$

$$-\int p(z \mid x)\mathrm{d}z + \int (\hat{x} - x)\frac{\partial p(z \mid x)}{\partial x}\mathrm{d}z = 0 \tag{14-48}$$

$$\int (\hat{x} - x)\frac{\partial p(z \mid x)}{\partial x}\mathrm{d}z = \int p(z \mid x)\mathrm{d}z \tag{14-49}$$

$$\int (\hat{x} - x)\frac{\partial p(z \mid x)}{\partial x}\mathrm{d}z = 1 \tag{14-50}$$

注意

$$\frac{\partial}{\partial x}\ln p(z \mid x) = \frac{1}{p(z \mid x)}\frac{\partial p(z \mid x)}{\partial x} \tag{14-51}$$

式(14-50)变为

$$\int (\hat{x} - x)p(z \mid x)\frac{\partial}{\partial x}\ln p(z \mid x)\mathrm{d}z = 1 \tag{14-52}$$

$$1 = \int \left[(\hat{x} - x)\sqrt{p(z \mid x)}\right]\left[\frac{\partial}{\partial x}\ln p(z \mid x)\sqrt{p(z \mid x)}\right]\mathrm{d}z \tag{14-53}$$

对等式两边求平方，得

$$1 = \left[\int[(\hat{x} - x)\ \sqrt{p(z \mid x)}]\left[\frac{\partial}{\partial x}\ln p(z \mid x)\ \sqrt{p(z \mid x)}\right]dz\right\}^2 \quad (14-54)$$

利用柯西不等式,得

$$1 \leqslant \int(\hat{x} - x)^2 p(z \mid x)dz\int\left[\frac{\partial}{\partial x}\ln p(z \mid x)\right]^2 p(z \mid x)dz \quad (14-55)$$

$$\int(\hat{x} - x)^2 p(z \mid x)dz \geqslant \frac{1}{\int\left[\frac{\partial}{\partial x}\ln p(z \mid x)\right]^2 p(z \mid x)dz} \quad (14-56)$$

式(14-56)的左边为变量 σ_x^2,而右边的分母可看作是 Fisher 信息 $J \triangleq \varepsilon\left\{\left[\frac{\partial}{\partial x}\ln p(z|x)\right]^2\right\}$。这样得

$$\sigma_x^2 \geqslant \frac{1}{\varepsilon\left\{\left[\frac{\partial}{\partial x}\ln p(z \mid x)\right]^2\right\}} \triangleq \frac{1}{J} \quad (14-57)$$

这里将一维 CRLB 定义为

$$\mathrm{CRLB} = \frac{1}{\varepsilon\left\{\left[\frac{\partial}{\partial x}\ln p(z \mid x)\right]^2\right\}} \quad (14-58)$$

因为在蒙特卡洛仿真中,可以用均方根误差估计得方差,这样 CRLB 的平方根就成为均方根误差的下界。

14.4.2 多维情况

对于多维情况,Fisher 信息变为 Fisher 信息矩阵 $\boldsymbol{J}_n$,而 CRLB 变成 $\boldsymbol{J}_n^{-1}$。如果在状态动态模型中包含过程噪声,Fisher 信息矩阵的逆就是 $\boldsymbol{P}_{n|n}^{xx}$的下界,即

$$\boldsymbol{P}_{n|n}^{xx} \triangleq \varepsilon\{[\boldsymbol{x}_n - \hat{\boldsymbol{x}}_{n|n}]\ [\boldsymbol{x}_n - \hat{\boldsymbol{x}}_{n|n}]^{\mathrm{T}}\} \geqslant \boldsymbol{J}_n^{-1} \quad (14-59)$$

将真实状态轨迹定义为 $\boldsymbol{x}_{1:n} \triangleq \{\boldsymbol{x}_0, \boldsymbol{x}_1, \cdots, \boldsymbol{x}_n\}$;状态估计初始化为 $\boldsymbol{x}_0$,在未有任何观测值之前,概率为 $p(\boldsymbol{x}_0)$;使用观测值 $z_{1:n} \triangleq \{z_0, z_1, \cdots, z_n\}$,状态 $\boldsymbol{x}_n$ 在时刻 t_n 的无偏估计由滤波器输出确定 $\hat{\boldsymbol{x}}_{n|n}$;轨迹 $\boldsymbol{x}_{1:n}$的 CRLB 由轨迹信息矩阵 $\boldsymbol{J}_{1:n}$的逆给出,有

$$\varepsilon\{[\boldsymbol{x}_{1:n} - \hat{\boldsymbol{x}}_{1:n|1:n}]\ [\boldsymbol{x}_{1:n} - \hat{\boldsymbol{x}}_{1:n|1:n}]^{\mathrm{T}}\} \geqslant \boldsymbol{J}_{1:n}^{-1} = \boldsymbol{CRLB}_{\mathrm{traj}} \quad (14-60)$$

将式(14-57)推广至多维情况,得 $\boldsymbol{J}_{1:n}$的第一种形式为

$$\boldsymbol{J}_{1:n} \triangleq \varepsilon\{[\boldsymbol{\nabla}_{\boldsymbol{x}_{1:n}}\ln p(\boldsymbol{x}_{1:n}, \boldsymbol{z}_{1:n})]\ [\boldsymbol{\nabla}_{\boldsymbol{x}_{1:n}}\ln p(\boldsymbol{x}_{1:n}, \boldsymbol{z}_{1:n})]^{\mathrm{T}}\} \quad (14-61)$$

注意式(14-51)可写成多维形式为

$$\nabla_{x_{1:n}}^{T}\ln p(\boldsymbol{x}_{1:n},\boldsymbol{z}_{1:n}) = \frac{1}{p(\boldsymbol{x}_{1:n},\boldsymbol{z}_{1:n})}\nabla_{x_{1:n}}^{T}p(\boldsymbol{x}_{1:n},\boldsymbol{z}_{1:n}) \tag{14-62}$$

将式(14－62)代入式(14－61),得 $\boldsymbol{J}_{1:n}$ 的第二种形式为

$$\boldsymbol{J}_{1:n} \triangleq -\varepsilon\{\nabla_{x_{1:n}}\nabla_{x_{1:n}}^{T}\ln p(\boldsymbol{x}_{1:n},\boldsymbol{z}_{1:n})\} \tag{14-63}$$

在这里,必须注意:

(1) 如果目标状态包含过程噪声 $\boldsymbol{v}_n$,是随机的,选取 $p(\boldsymbol{x}_{1:n},z_{1:n})$;

(2) 如果目标状态是确定性的(零过程噪声),选取 $p(z_{1:n}|\boldsymbol{x}_{1:n})$;

(3) $\nabla_{x_{1:n}}$ 表示 $\boldsymbol{x}_{1:n}$ 计算的真实值。

14.4.3 递归法求解 CRLB

在 $\boldsymbol{J}_x$ 和 $\boldsymbol{J}_{1:n-1}$ 已知的情况下,求解 $\boldsymbol{J}_{1:n}$ 非常有意义,Tichavsky 等[2,4]就提出了这样一种方法。注意,可以将 $\boldsymbol{x}_{1:n}$ 和 $\nabla_{x_{1:n}}$ 分解为

$$\boldsymbol{x}_{1:n} = \begin{bmatrix} \boldsymbol{x}_{1:n-1} \\ \boldsymbol{x}_n \end{bmatrix} \tag{14-64}$$

$$\nabla_{x_{1:n}} = \begin{bmatrix} \nabla_{x_{1:n-1}} \\ \nabla_{x_n} \end{bmatrix} \tag{14-65}$$

这样 $\nabla_{x_{1:n}}\nabla_{x_{1:n}}^{T}$ 可分解为

$$\nabla_{x_{1:n}}\nabla_{x_{1:n}}^{T} = \begin{bmatrix} \nabla_{x_{1:n-1}} \\ \nabla_{x_n} \end{bmatrix}\begin{bmatrix} \nabla_{x_{1:n-1}}^{T} & \nabla_{x_n}^{T} \end{bmatrix} = \begin{bmatrix} \nabla_{x_{1:n-1}}\nabla_{x_{1:n-1}}^{T} & \nabla_{x_{1:n-1}}\nabla_{x_n}^{T} \\ \nabla_{x_n}\nabla_{x_{1:n-1}}^{T} & \nabla_{x_n}\nabla_{x_n}^{T} \end{bmatrix} \tag{14-66}$$

将式(14－66)代入式(14－63),可实现 $\boldsymbol{J}_{1:n}$ 的分解,即

$$\boldsymbol{J}_{1:n} = \begin{bmatrix} \boldsymbol{A}_{1:n-1|1:n-1} & \boldsymbol{B}_{1:n-1|n} \\ \boldsymbol{B}_{1:n-1|n}^{T} & \boldsymbol{C}_{n|n} \end{bmatrix} \tag{14-67}$$

$$\boldsymbol{A}_{1:n-1|1:n-1} \triangleq -\varepsilon\{\nabla_{x_{1:n-1}}\nabla_{x_{1:n-1}}^{T}\ln p(\boldsymbol{x}_{1:n},\boldsymbol{z}_{1:n})\} \tag{14-68}$$

$$\boldsymbol{B}_{1:n-1|n} \triangleq -\varepsilon\{\nabla_{x_{1:n-1}}\nabla_{x_n}^{T}\ln p(\boldsymbol{x}_{1:n},\boldsymbol{z}_{1:n})\} \tag{14-69}$$

$$\boldsymbol{C}_{n|n} \triangleq -\varepsilon\{\nabla_{x_n}\nabla_{x_n}^{T}\ln p(\boldsymbol{x}_{1:n},\boldsymbol{z}_{1:n})\} \tag{14-70}$$

利用式(2－19)求得分块矩阵(14－67)的逆,求得 $\boldsymbol{J}_{1:n}^{-1}$,即

$$\boldsymbol{J}_{1:n}^{-1} = \begin{bmatrix} \boldsymbol{T}_{11} & \boldsymbol{T}_{12} \\ \boldsymbol{T}_{21} & \boldsymbol{T}_{22} \end{bmatrix} \tag{14-71}$$

$$T_{11} \triangleq (A_{1:n-1|1:n-1} - B_{1:n-1|n} C_{n|n}^{-1} B_{1:n-1|n}^{\mathrm{T}})^{-1} \tag{14-72}$$

$$T_{12} \triangleq -A_{1:n-1|1:n-1} B_{1:n-1|n}^{-1} \cdot (C_{n|n} - B_{1:n-1|n}^{\mathrm{T}} A_{1:n-1|1:n-1}^{-1} B_{1:n-1|n})^{-1} \tag{14-73}$$

$$T_{21} \triangleq -C_{n|n}^{-1} B_{1:n-1|n}^{\mathrm{T}} (A_{1:n-1|1:n-1} - B_{1:n-1|n} C_{n|n}^{-1} B_{1:n-1|n}^{\mathrm{T}})^{-1} \tag{14-74}$$

$$T_{22} \triangleq (C_{n|n} - B_{1:n-1|n}^{\mathrm{T}} A_{1:n-1|1:n-1}^{-1} B_{1:n-1|n})^{-1} \tag{14-75}$$

矩阵$J_{1:n}^{-1}$中右下角的块可记为J_n^{-1},为$\hat{x}_{n|n}$的 Cramer – Rao 下界,即

$$J_n^{-1} = (C_{n|n} - B_{1:n-1|n}^{\mathrm{T}} A_{1:n-1|1:n-1}^{-1} B_{1:n-1|n})^{-1} \tag{14-76}$$

因此有

$$J_n = C_{n|n} - B_{1:n-1|n}^{\mathrm{T}} A_{1:n-1|1:n-1}^{-1} B_{1:n-1|n} \tag{14-77}$$

注意,式(14 –76)需要对一个大型矩阵$A_{1:n-1|1:n-1}$进行逆运算。为解决这个问题,本书提出一种递归关系,由 J_n 得 J_{n-1}而不需要对大矩阵求逆。考虑式(14 –61)中的联合密度,有

$$\begin{aligned} p(x_{1:n}, z_{1:n}) &= p(x_n, z_n, x_{1:n-1}, z_{1:n-1}) \\ &= p(z_n \mid x_n, x_{1:n-1}, z_{1:n-1}) p(x_n \mid x_{1:n-1}, z_{1:n-1}) \cdot \\ &\quad p(x_{1:n-1}, z_{1:n-1}) \end{aligned} \tag{14-78}$$

由于z_n 只取决于x_n,而x_n 只取决于x_{n-1},式(14 – 78)变为

$$p(x_{1:n}, z_{1:n}) = p(z_n \mid x_n) p(x_n \mid x_{n-1}) p(x_{1:n-1}, z_{1:n-1}) \tag{14-79}$$

若将$\nabla_{x_{1:n}}$分解为

$$\begin{aligned} \nabla_{x_{1:n}} \nabla_{x_{1:n}}^{\mathrm{T}} &= \begin{bmatrix} \nabla_{x_{1:n}} - 2 \\ \nabla_{x_{n-1}} \\ \nabla_{x_n} \end{bmatrix} [\nabla_{x_{1:n-2}}^{\mathrm{T}} \quad \nabla_{x_{n-1}}^{\mathrm{T}} \quad \nabla_{x_n}^{\mathrm{T}}] \\ &= \begin{bmatrix} \nabla_{x_{1:n-2}} \nabla_{x_{1:n-2}}^{\mathrm{T}} & \nabla_{x_{1:n-2}} \nabla_{x_{n-1}}^{\mathrm{T}} & \nabla_{x_{1:n-2}} \nabla_{x_n}^{\mathrm{T}} \\ \nabla_{x_{n-1}} \nabla_{x_{1:n-2}}^{\mathrm{T}} & \nabla_{x_{n-1}} \nabla_{x_{n-1}}^{\mathrm{T}} & \nabla_{x_{n-1}} \nabla_{x_n}^{\mathrm{T}} \\ \nabla_{x_n} \nabla_{x_{1:n-2}}^{\mathrm{T}} & \nabla_{x_n} \nabla_{x_{n-1}}^{\mathrm{T}} & \nabla_{x_n} \nabla_{x_n}^{\mathrm{T}} \end{bmatrix} \end{aligned} \tag{14-80}$$

则轨迹信息矩阵(14 –63)变成

$$J_{1:n} = \begin{bmatrix} S_{11} & S_{12} & S_{13} \\ S_{21} & S_{22} & S_{23} \\ S_{31} & S_{32} & S_{33} \end{bmatrix} \tag{14-81}$$

这里

$$\begin{aligned}\boldsymbol{S}_{11} &= -\varepsilon\{\boldsymbol{\nabla}_{\boldsymbol{x}_{1:n-2}}[\boldsymbol{\nabla}^{\mathrm{T}}_{\boldsymbol{x}_{1:n-2}}\ln p(\boldsymbol{x}_{1:n},\boldsymbol{z}_{1:n})]\} \\ &= -\varepsilon\{\boldsymbol{\nabla}_{\boldsymbol{x}_{1:n-2}}[\boldsymbol{\nabla}^{\mathrm{T}}_{\boldsymbol{x}_{1:n-2}}[\ln p(\boldsymbol{x}_{1:n-1},\boldsymbol{z}_{1:n-1}) + \\ &\quad \ln p(z_n \mid \boldsymbol{x}_n) + \ln p(\boldsymbol{x}_n \mid \boldsymbol{x}_{n-1})]]\} \\ &= -\varepsilon\{\boldsymbol{\nabla}_{\boldsymbol{x}_{1:n-2}}[\boldsymbol{\nabla}^{\mathrm{T}}_{\boldsymbol{x}_{1:n-2}}\ln p(\boldsymbol{x}_{1:n-1},\boldsymbol{z}_{1:n-1})]\} \\ &= \boldsymbol{A}_{1:n-2|1:n-2}\end{aligned} \tag{14-82}$$

同样有

$$\boldsymbol{S}_{22} = \boldsymbol{C}_{n-1|n-1} + \boldsymbol{D}^{11}_{n-1|n-1} \tag{14-83}$$

$$\boldsymbol{S}_{12} = \boldsymbol{B}_{1:n-2|n-1} \tag{14-84}$$

$$\boldsymbol{S}_{31} = \boldsymbol{0} \tag{14-85}$$

$$\boldsymbol{S}_{32} = \boldsymbol{D}^{12}_{n|n-1} \tag{14-86}$$

$$\boldsymbol{S}_{22} = \boldsymbol{D}^{22}_{n|n} \tag{14-87}$$

其中 $\boldsymbol{A}_{1:n-2|1:n-2}$、$\boldsymbol{B}_{1:n-2|n-1}$ 和 $\boldsymbol{C}_{n-1|n-1}$ 分别由式(14－68)、式(14－69)和式(14－70)定义,而

$$\boldsymbol{D}^{11}_{n-1|n-1} \triangleq -\varepsilon\{\boldsymbol{\nabla}_{\boldsymbol{x}_{n-1}}\boldsymbol{\nabla}^{\mathrm{T}}_{\boldsymbol{x}_{n-1}}\ln p(\boldsymbol{x}_n \mid \boldsymbol{x}_{n-1})\} \tag{14-88}$$

$$\boldsymbol{D}^{21}_{n-1|n} \triangleq -\varepsilon\{\boldsymbol{\nabla}_{\boldsymbol{x}_{n-1}}\boldsymbol{\nabla}^{\mathrm{T}}_{\boldsymbol{x}_n}\ln p(\boldsymbol{x}_n \mid \boldsymbol{x}_{n-1})\} \tag{14-89}$$

$$\boldsymbol{D}^{12}_{n|n-1} = [\boldsymbol{D}^{21}_{n-1|n}]^{\mathrm{T}} \tag{14-90}$$

$$\boldsymbol{D}^{22}_{n|n} \triangleq -\varepsilon\{\boldsymbol{\nabla}_{\boldsymbol{x}_n}\boldsymbol{\nabla}^{\mathrm{T}}_{\boldsymbol{x}_n}\ln p(\boldsymbol{x}_n \mid \boldsymbol{x}_{n-1})\} - \varepsilon\{\boldsymbol{\nabla}_{\boldsymbol{x}_n}\boldsymbol{\nabla}^{\mathrm{T}}_{\boldsymbol{x}_n}\ln p(z_n \mid \boldsymbol{x}_n)\} \tag{14-91}$$

可将式(14－81)写为

$$\boldsymbol{J}_{1:n} = \begin{bmatrix} \boldsymbol{A}_{1:n-2|1:n-2} & \boldsymbol{B}_{1:n-2|n-1} & \boldsymbol{0} \\ \boldsymbol{B}^{\mathrm{T}}_{1:n-2|n-1} & \boldsymbol{C}_{n-1|n-1} + \boldsymbol{D}^{11}_{n-1|n-1} & \boldsymbol{D}^{12}_{n|n-1} \\ \boldsymbol{0} & \boldsymbol{D}^{21}_{n-1|n} & \boldsymbol{D}^{22}_{n|n} \end{bmatrix} \tag{14-92}$$

用式(2－19)求得分块矩阵(14－92)的逆,计算 $\boldsymbol{J}^{-1}_{1:n}$,这里只取右下角块的逆,有

$$\begin{aligned}\boldsymbol{J}_n &= \boldsymbol{D}^{22}_{n|n} - [\boldsymbol{0} \quad \boldsymbol{D}^{21}_{n-1|n}] \cdot \begin{bmatrix} \boldsymbol{A}_{1:n-2|1:n-2} & \boldsymbol{B}_{1:n-2|n-1} \\ \boldsymbol{B}^{\mathrm{T}}_{1:n-2|n-1} & \boldsymbol{C}_{n-1|n-1} + \boldsymbol{D}^{11}_{n-1|n-1} \end{bmatrix} \begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{D}^{12}_{n|n-1} \end{bmatrix} \\ &= \boldsymbol{D}^{22}_{n|n} - \boldsymbol{D}^{21}_{n-1|n}[\boldsymbol{C}_{n-1|n-1} + \boldsymbol{D}^{11}_{n-1|n-1} - \boldsymbol{B}^{\mathrm{T}}_{1:n-2|n-1}\boldsymbol{A}^{-1}_{1:n-2|1:n-2}\boldsymbol{B}_{1:n-2|n-1}]^{-1} \cdot \boldsymbol{D}^{12}_{n|n-1}\end{aligned} \tag{14-93}$$

代入式(14－77),则式(14－93)变成 $\boldsymbol{J}_n$ 的递归,即

$$\boldsymbol{J}_n = \boldsymbol{D}^{22}_{n|n} - \boldsymbol{D}^{21}_{n-1|n}[\boldsymbol{J}_{n-1} + \boldsymbol{D}^{11}_{n-1|n-1}]^{-1}\boldsymbol{D}^{12}_{n|n-1} \tag{14-94}$$

用式(14-61)初始化这一递归,设

$$\boldsymbol{J}_0 = \varepsilon\{[\boldsymbol{\nabla}_{\boldsymbol{x}_0}\ln p(\boldsymbol{x}_0)][\boldsymbol{\nabla}_{\boldsymbol{x}_0}\ln p(\boldsymbol{x}_0)]^{\mathrm{T}}\} \tag{14-95}$$

14.4.4 高斯加性噪声的 Cramer - Rao 下界

考虑动态模型和观测模型均为高斯分布的情况,即

$$\boldsymbol{x}_n = \boldsymbol{f}_{n-1}(\boldsymbol{x}_n) + \boldsymbol{v}_{n-1} \tag{14-96}$$

$$\boldsymbol{z}_n = \boldsymbol{h}_n(\boldsymbol{x}_n) + \boldsymbol{w}_{n-1} \tag{14-97}$$

$$\boldsymbol{v}_n \sim \mathcal{N}(\boldsymbol{0},\boldsymbol{Q}) \tag{14-98}$$

$$\boldsymbol{w}_n \sim \mathcal{N}(\boldsymbol{0},\boldsymbol{R}) \tag{14-99}$$

如果初始分布是高斯的,即

$$p(\boldsymbol{x}_0) = \mathcal{N}(\boldsymbol{x}_0;\hat{\boldsymbol{x}}_0,\boldsymbol{P}_0^{xx}) \tag{14-100}$$

那么则有

$$\begin{aligned}\boldsymbol{\nabla}_{\boldsymbol{x}_0}\ln p(\boldsymbol{x}_0) &= \boldsymbol{\nabla}_{\boldsymbol{x}_0}\left\{c - \frac{1}{2}[(\boldsymbol{x}_0 - \hat{\boldsymbol{x}}_0)^{\mathrm{T}}[\boldsymbol{P}_0^{xx}]^{-1}(\boldsymbol{x}_0 - \hat{\boldsymbol{x}}_0)]\right\}\\ &= -[\boldsymbol{P}_0^{xx}]^{-1}(\boldsymbol{x}_0 - \hat{\boldsymbol{x}}_0)\end{aligned} \tag{14-101}$$

这样,由式(14-95)以及$[\boldsymbol{P}_0^{xx}]^{-\mathrm{T}} = [\boldsymbol{P}_0^{xx}]^{-1}$,$\boldsymbol{J}_0$ 变成

$$\begin{aligned}\boldsymbol{J}_0 &= \varepsilon\{[\boldsymbol{P}_0^{xx}]^{-1}(\boldsymbol{x}_0 - \hat{\boldsymbol{x}}_0)(\boldsymbol{x}_0 - \hat{\boldsymbol{x}}_0)^{\mathrm{T}}[\boldsymbol{P}_0^{xx}]^{-1}\}\\ &= [\boldsymbol{P}_0^{xx}]^{-1}\varepsilon\{(\boldsymbol{x}_0 - \hat{\boldsymbol{x}}_0)(\boldsymbol{x}_0 - \hat{\boldsymbol{x}}_0)^{\mathrm{T}}\}[\boldsymbol{P}_0^{xx}]^{-1} = [\boldsymbol{P}_0^{xx}]^{-1}\end{aligned} \tag{14-102}$$

类似地,有

$$\boldsymbol{\nabla}_{\boldsymbol{x}_{n-1}}\ln p(\boldsymbol{x}_n \mid \boldsymbol{x}_{n-1}) = [\boldsymbol{\nabla}_{\boldsymbol{x}_{n-1}}\boldsymbol{f}_{n-1}^{\mathrm{T}}(\boldsymbol{x}_{n-1})]\boldsymbol{Q}_{n-1}^{-1}\boldsymbol{v}_{n-1} \tag{14-103}$$

$$\boldsymbol{\nabla}_{\boldsymbol{x}_n}\ln p(\boldsymbol{z}_n \mid \boldsymbol{x}_n) = [\boldsymbol{\nabla}_{\boldsymbol{x}_n}\boldsymbol{h}_n^{\mathrm{T}}(\boldsymbol{x}_n)]\boldsymbol{R}_n^{-1}\boldsymbol{w}_n \tag{14-104}$$

这样式(14-88)-式(14-91)变成

$$\boldsymbol{D}_{n-1|n-1}^{11} = \varepsilon\{\widetilde{\boldsymbol{F}}_{n-1}^{\mathrm{T}}\boldsymbol{Q}_{n-1}^{-1}\widetilde{\boldsymbol{F}}_{n-1}\} \tag{14-105}$$

$$\boldsymbol{D}_{n-1|n}^{12} = -\varepsilon\{\widetilde{\boldsymbol{F}}_{n-1}^{\mathrm{T}}\}\boldsymbol{Q}_{n-1}^{-1} \tag{14-106}$$

$$\boldsymbol{D}_{n|n-1}^{21} = -\boldsymbol{Q}_{n-1}^{-1}\varepsilon\{\widetilde{\boldsymbol{F}}_{n-1}\} \tag{14-107}$$

$$\boldsymbol{D}_{n|n}^{22} = \boldsymbol{Q}_{n-1}^{-1} + \varepsilon\{\widetilde{\boldsymbol{H}}_n^{\mathrm{T}}\boldsymbol{R}_n^{-1}\widetilde{\boldsymbol{H}}_n\} \tag{14-108}$$

$$\widetilde{\boldsymbol{F}}_{n-1} \triangleq [\boldsymbol{\nabla}_{\boldsymbol{x}_{n-1}}\boldsymbol{f}_{n-1}^{\mathrm{T}}(\boldsymbol{x}_{n-1})]_{\boldsymbol{x}_{n-1}(\mathrm{TRUE})}^{\mathrm{T}} \tag{14-109}$$

$$\widetilde{\boldsymbol{H}}_n \triangleq [\boldsymbol{\nabla}_{\boldsymbol{x}_n}\boldsymbol{h}_n^{\mathrm{T}}(\boldsymbol{x}_n)]_{\boldsymbol{x}_n(\mathrm{TRUE})}^{\mathrm{T}} \tag{14-110}$$

总之,对于高斯噪声下的非线性模型,可以由式(14-95)、式(14-94),并用式(14-105)~式(14-108)代替式(14-88)~式(14-91),递归计算CRLB。

在CRLB计算中会遇到两个问题。第一个是式(14-105)~式(14-108)中$\varepsilon\{\cdot\}$的计算。在高斯分布情况下,因为$\widetilde{\boldsymbol{F}}_{n-1}$和$\widetilde{\boldsymbol{H}}_n$分别在轨迹点$\boldsymbol{x}_{n-1}$和$\boldsymbol{x}_n$进行计算,可以用蒙特卡洛方法来估计这些期望值。利用式(14-96),噪声可由$\boldsymbol{v}_n \sim \mathcal{N}(\boldsymbol{0},\boldsymbol{Q})$采样得到,可生成$M$组蒙特卡洛轨迹值$\{\boldsymbol{x}_{m,n};n=1,\cdots,N;m=1,\cdots,M\}$。对每个时间点的蒙特卡洛轨迹求平均,从而得期望值的估计。第二个问题是$\boldsymbol{Q}_n$当为奇异时出现,此时它的逆不存在。对这种特殊情况的处理方法超出了本书的研究范畴,感兴趣的读者可以参阅文献[4](也可以在文献[2]的第686页找到)。

14.4.5 零过程噪声的高斯 Cramer-Rao 下界

当过程噪声非常小的时候,可以假设过程噪声为零,以计算CRLB。当式(14-96)中过程噪声为零,状态矢量的演化就是完全确定的。因此,可以忽略式(14-105)~式(14-108)中的期望运算。这样式(14-94)变成

$$\boldsymbol{J}_n = \boldsymbol{Q}_{n-1}^{-1} + \widetilde{\boldsymbol{H}}_n^{\mathrm{T}}\boldsymbol{R}_n^{-1}\widetilde{\boldsymbol{H}}_n - \boldsymbol{Q}_{n-1}^{-1}\widetilde{\boldsymbol{F}}_{n-1}[\boldsymbol{J}_{n-1} + \widetilde{\boldsymbol{F}}_{n-1}^{\mathrm{T}}\boldsymbol{Q}_{n-1}^{-1}\widetilde{\boldsymbol{F}}_{n-1}]^{-1}\widetilde{\boldsymbol{F}}_{n-1}^{\mathrm{T}}\boldsymbol{Q}_{n-1}^{-1} \tag{14-111}$$

使用式(2-13)的矩阵求逆引理,上式可简化为

$$\boldsymbol{J}_n = (\boldsymbol{Q}_{n-1} + \widetilde{\boldsymbol{F}}_{n-1}\boldsymbol{J}_{n-1}^{-1}\widetilde{\boldsymbol{F}}_{n-1}^{\mathrm{T}})^{-1} + \widetilde{\boldsymbol{H}}_n^{\mathrm{T}}\boldsymbol{R}_n^{-1}\widetilde{\boldsymbol{H}}_n \tag{14-112}$$

因为过程噪声为零,有

$$\boldsymbol{J}_n = (\widetilde{\boldsymbol{F}}_{n-1}^{-1})^{\mathrm{T}}\boldsymbol{J}_{n-1}\widetilde{\boldsymbol{F}}_{n-1}^{-1} + \widetilde{\boldsymbol{H}}_n^{\mathrm{T}}\boldsymbol{R}_n^{-1}\widetilde{\boldsymbol{H}}_n \tag{14-113}$$

通过$\boldsymbol{J}_n^{-1}$的逆并应用矩阵求逆引理,式(14-113)简化为与卡尔曼滤波器协方差更新方程式(5-30)相同的形式,需计算真值$\boldsymbol{x}_n$的雅可比。这样,当没有过程噪声时,非线性滤波器的CRLB递归与EKF的协方差更新方程形式相同,需要计算真实状态矢量$\boldsymbol{x}_n$的雅可比。

14.4.6 线性模型的高斯 Cramer-Rao 下界

考虑下面的情况,用线性方程代替式(14-96)和式(14-97)

$$\boldsymbol{x}_n = \boldsymbol{F}\boldsymbol{x}_n + \boldsymbol{v}_{n-1} \tag{14-114}$$

$$\boldsymbol{z}_n = \boldsymbol{H}\boldsymbol{x}_n + \boldsymbol{w}_{n-1} \tag{14-115}$$

式(14-109)和式(14-110)的雅可比变成

$$\widehat{F}_{n-1} \triangleq [\nabla_{x_{n-1}} x_n^{\mathrm{T}} F^{\mathrm{T}}]^{\mathrm{T}}_{x_{n-1}(\mathrm{TRUE})} = F \tag{14-116}$$

$$\widehat{H}_n \triangleq [\nabla_{x_n} x_n^{\mathrm{T}} H^{\mathrm{T}}]^{\mathrm{T}}_{x_n(\mathrm{TRUE})} = H \tag{14-117}$$

因此，对于线性高斯情况，雅可比独立于目标状态，可以消去式（14－105）~式（14－108）中的期望值，得

$$D^{11}_{n-1|n-1} = F^{\mathrm{T}} Q^{-1}_{n-1} F \tag{14-118}$$

$$D^{12}_{n-1|n-1} = -F^{\mathrm{T}} Q^{-1}_{n-1} \tag{14-119}$$

$$D^{21}_{n-1|n-1} = -Q^{-1}_{n-1} F \tag{14-120}$$

$$D^{11}_{n-1|n-1} = Q^{-1}_{n-1} + H^{\mathrm{T}} R^{-1}_n H \tag{14-121}$$

把这些代入式（14－94），利用第2.1.4的分块矩阵求逆引理，将式（14－94）简化为

$$J^{-1}_n = P^{xx}_{n|n} \tag{14-122}$$

因此，在高斯线性情况下，CRLB可简化为线性卡尔曼滤波器的更新状态协方差矩阵，这说明在这种情况下，线性卡尔曼滤波器是最优的。

14.5 卡尔曼类DIFAR轨迹估计器的性能

DIFAR案例研究中用到的跟踪估计器包括EKF、UKF、SSKF、GHKF和MCKF。本节比较了每种跟踪估计器在DIFAR跟踪中x轴和y轴方向上的均方根误差。图14－3~图14－7给出信号信噪比从20dB降至0dB、步长为5dB的情况下，5种不同跟踪算法在x轴和y轴方向上的均方根误差及CRLB。如图所示，当应用于DIFAR问题时，所有信噪比下，不同跟踪算法之间只有很小的差别。

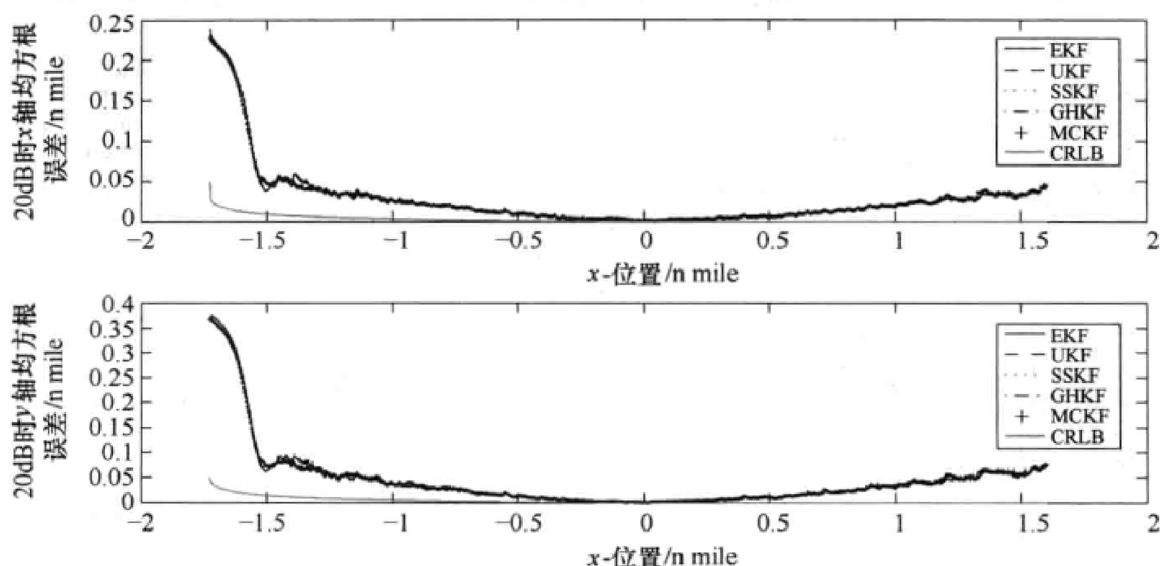


图14－3　当输入信号信噪比为20dB时，5种不同跟踪估计算法均方根误差比较

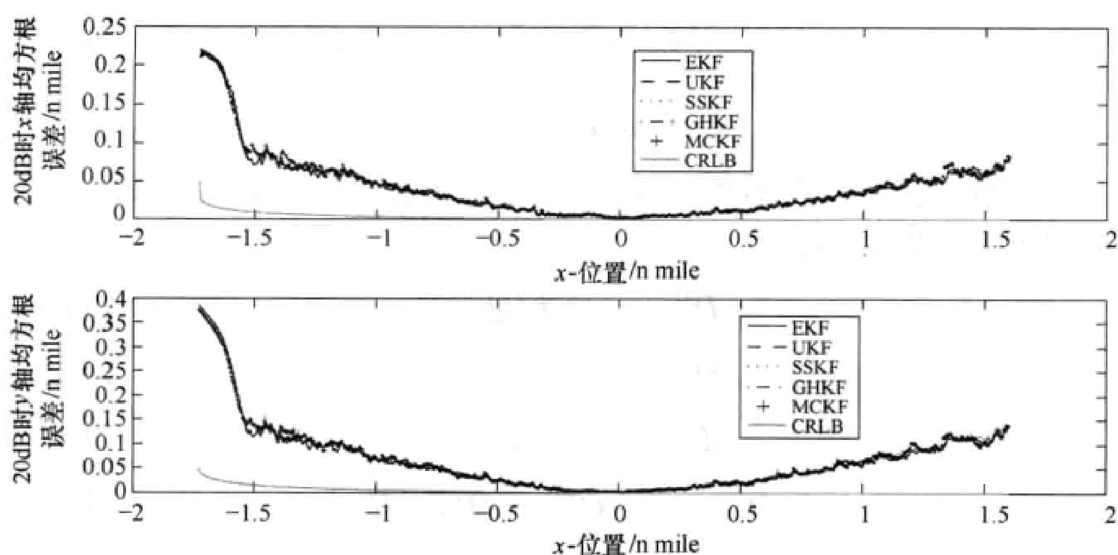


图 14－4　当输入信号信噪比为 15dB 时,5 种不同跟踪估计算法均方根误差比较

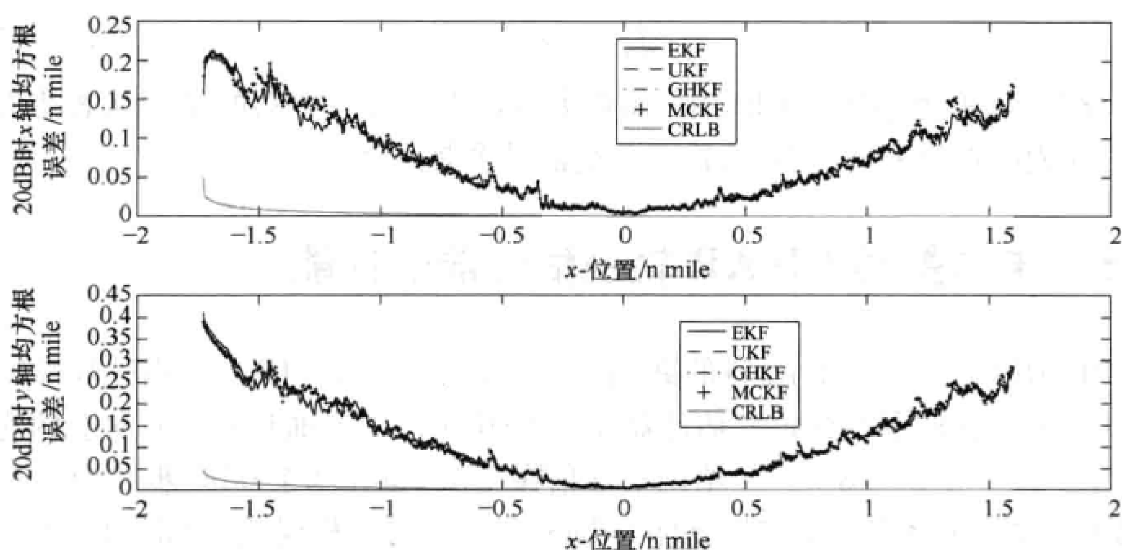


图 14－5　当输入信号信噪比为 10dB 时,5 种不同跟踪估计算法均方根误差比较

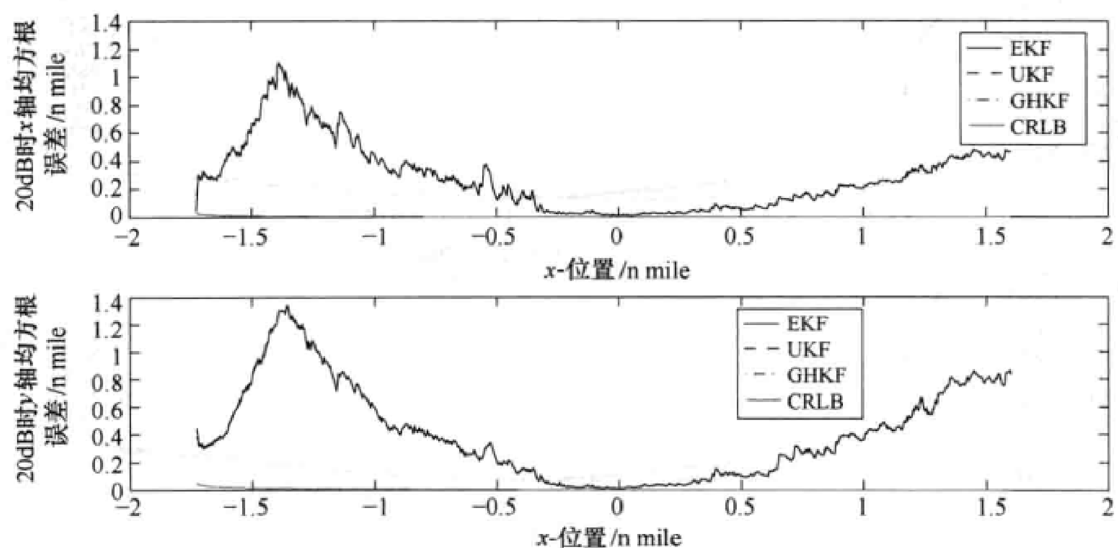


图 14－6　当输入信号信噪比为 5dB 时,5 种不同跟踪估计算法均方根误差比较

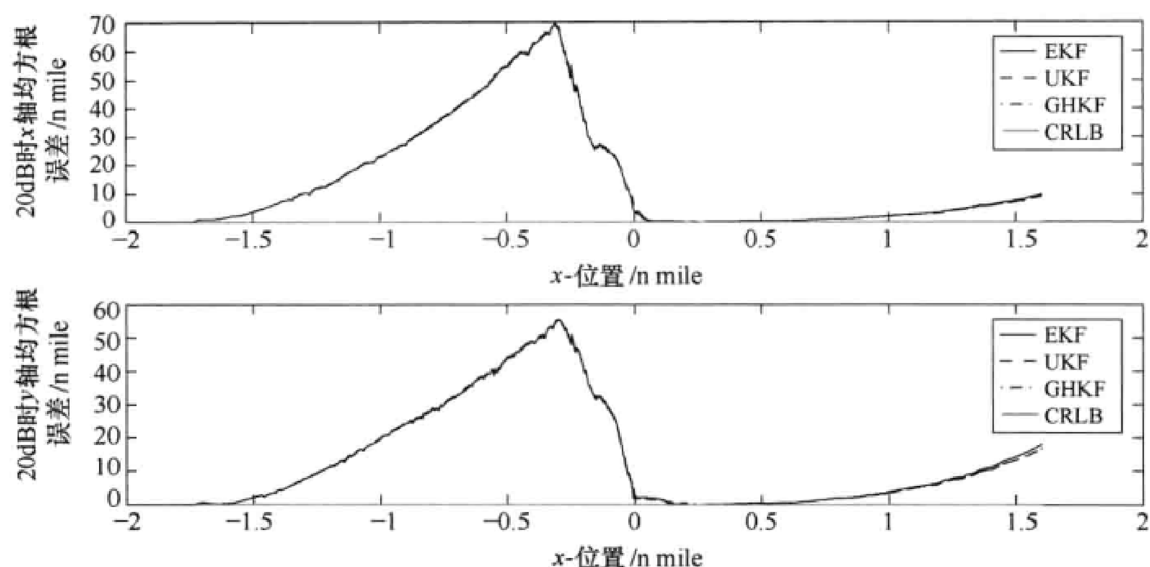


图 14－7　当输入信号信噪比为 0dB 时,5 种不同跟踪估计算法均方根误差比较

每幅图中下端的灰色线条为 CRLB。基于这些性能的度量可知,对于 DIFAR 问题,应该采用卡尔曼滤波器,因为它能最大限度地减少计算载荷。从这些数据很容易看到,对于所有信噪比的输入信号,当目标船舶进入浮标场接近(0,0)的位置时,滤波器的跟踪性能最好。但是当目标离浮标场很远时,随着与浮标场中心距离的增加以及信噪比降低,跟踪性能降幅增加。此外,信噪比低于 5dB 时,SSKF 跟踪器变得不稳定;当信噪比为 0dB 时,MCKF 也变得不稳定。在 RMS 图中不包括这些情况。

参考文献

[1] Van Trees HL. Detection, Estimation, and Modulation theory, Part I: Detection, Estimation, and Linear Modulation Theory. Wiley, 1968.

[2] Van Trees HL. Bell KL. Bayesian Bounds for Parameter Estimation and Nonlinear Filtering/Tracking. Wiley Interscience, 2007.

[3] Ristic B, Arulampa$_1$am S. Gordon N. Beyond the Kalman Filter: Particle Filters for Tracking Applications. Artech House, 2004.

[4] Tichavsky P, Muavchik CH, Nehorai A. Posterior Cramer－Rao bounds for discrete－time nonlinear filtering. *IEEE Trans. Signal Process*, 1998, 46(5): 1386－1396.

BAYESIAN

第Ⅲ部分

蒙特卡洛方法

第15章 蒙特卡洛方法概述

第Ⅱ部分研究了计算非线性函数高斯加权矩积分的数值方法。其中多数方法,都采用多维多项式展开逼近非线性函数,用确定性矢量点集合计算高斯矩积分,得 Sigma 点卡尔曼滤波器。第 12 章中给出了另一种方法,用一组离散的随机样本来逼近多维高斯密度,而不是逼近非线性函数。该方法可以用非线性函数在采样点处的加权和计算矩积分。当采样点的数目增加时,逼近变得更为精确。蒙特卡洛采样方法可以用来直接求解一些具有解析形式的密度函数,例如学生 t 密度、瑞利密度和卡方密度,但首先要考虑的问题是这种蒙特卡洛积分方法是否适用于密度未知或难于采样的情况。

本章首先研究如何仅仅基于一组密度的随机采样,来估算密度的形式。那么如何获得最初的采样呢? 这里提出的方法包括多维直方图的生成。

之后提出了直方图的核密度逼近,以提供更平滑的密度函数估计。当应用于多峰密度或高度非高斯密度的可视化、以及在未提供任何有意义的信息情况下的矩估计时,核估计的实用性变得非常明显。

Richardson 等[1]给出实例,密度在两个维度上是多峰的,矩不代表任何模式的峰值(通常会有这样一个错误的印象,峰值可能对应感兴趣的目标的位置)。在下一章讨论重采样方法时,还将发现核密度估计的作用。在讨论完核密度估计方法之后,继续回到如何由未知密度通过引入重要性采样概念来生成样本。最后一节对本章的内容进行了概括与总结。

15.1 通过蒙特卡洛样本集估计密度

15.1.1 由二维高斯混合密度生成样本

本节将讨论一些仅仅基于密度的采样,来估算密度形式的方法。通常从密

度的采样(数据)来推断未知密度,因此必须采用一种非参数方法。这种非参数密度估计方法可以提供视觉线索,揭示偏态分布或数据中存在多个模式,可以进行观测值的重要物理解释[2]。例如,考虑双模高斯混合密度

$$p(\boldsymbol{x}) = w_1 \mathcal{N}(\boldsymbol{x}_1;\hat{\boldsymbol{x}}_1,\boldsymbol{\Sigma}_1) + w_2 \mathcal{N}(\boldsymbol{x}_2;\hat{\boldsymbol{x}}_2,\boldsymbol{\Sigma}_2) \tag{15-1}$$

这里 $\omega_1 + \omega_2 = 1$。由这样一个混合密度可以很容易进行采样,如果所需的样本总数为 N,从$\mathcal{N}(\boldsymbol{x}_1;\hat{\boldsymbol{x}}_1,\boldsymbol{\Sigma}_1)$中选取 $\omega_1 N$ 个采样,从$\mathcal{N}(\boldsymbol{x}_2,\hat{\boldsymbol{x}}_2,\boldsymbol{\Sigma}_2)$中选取 $\omega_2 N$ 个采样。图 15-1 中 $N=1000$ 且$(\omega_1,\omega_2)=(0.3,0.7)$。同时,设置

$$\hat{\boldsymbol{x}}_1 = (-5, -8)$$

$$\hat{\boldsymbol{x}}_2 = (7,7)$$

$$\boldsymbol{\Sigma}_1 = \begin{bmatrix} 100 & -5 \\ -5 & 1 \end{bmatrix}$$

$$\boldsymbol{\Sigma}_2 = \begin{bmatrix} 4 & 8 \\ 8 & 64 \end{bmatrix}$$

15.1.2 利用多维直方图拟合密度

本节提出一种可视化的概率密度函数,注意落在给定区域的样本概率由落在该区域的样本的归一化密度得。这样可将归一化的二维直方图作为首选的估计和可视化方法,因为它实际上度量的是特定的二维区间内观测到样本的概率。可将多维直方图密度估计定义为

$$\hat{p}(\boldsymbol{x}) = \frac{\nu_{ik}}{NA_{ik}} \tag{15-2}$$

式中:ν_{ik}为落在超体积条(bin)A_{ik}内的样本数目。例如,对于一个二维密度$A_{ik} = (x_{i+1}-x_i)(y_{k+1}-y_k)$,如果所有的条大小相同,可以令 $h=(x_{i+1}-x_i)=(y_{k+1}-y_k)$, $\forall i,k$ 且

$$\hat{p}(\boldsymbol{x}) = \frac{\nu_{ik}}{Nh^2} \tag{15-3}$$

这样,对于一般 n_x-维直方图,有

$$\hat{p}(\boldsymbol{x}) = \frac{\nu_i}{Nh^{n_x}} \tag{15-4}$$

式中:ν_i 为落在第 i 个超体积条内的样本数目。后面的直方图中只有一个自由参数——条的宽度 h。直方图可以渐近逼近任意连续密度,因此可以称之为非参数估计。带宽 h 的选取是近期研究的重点,此处不做详细介绍,可参阅 Scott[3] 及其包含的参考文献。Scott 给出如何使用均方误差求和(IMSE)以及误差平方和的均值(MISE),基于样本数据估计最优带宽。

可以从 MATLAB 中心[4]下载生成并绘制二维直方图的 MATLAB 程序，注意，这些子程序必须适合特定的数据集，在编写时是非规范化的。图 15－2 中给出了由图 15－1 中数据生成的二维规范化直方图。根据数据的最大范围，在每个维度上选取合适带宽，产生 99 个条。

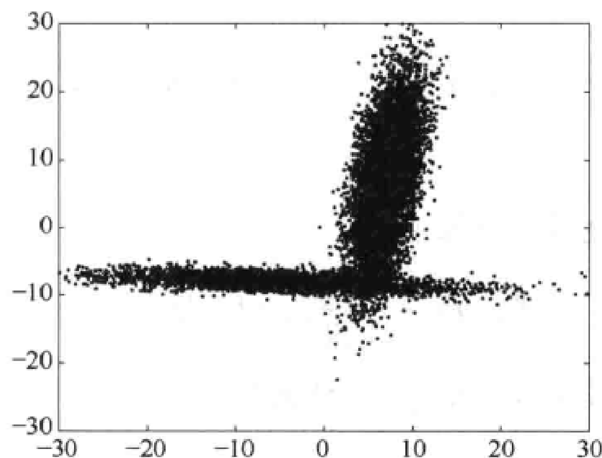


图 15－1　二维高斯混合密度的采样

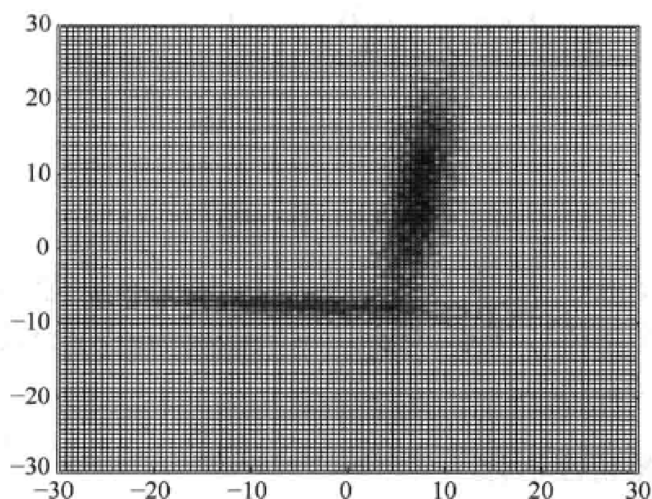


图 15－2　基于高斯混合分布采样的二维直方图

15.1.3　核密度估算

式(15－3)给出了落在中心为 $\boldsymbol{x}\in\mathbf{R}^2$ 的给定二维条内样本的概率估计。在条区域为零的极限情况下，每一个条中最多包含一个样本。这时，可以将样本密度函数写成

$$p_s(\boldsymbol{x}) = \frac{1}{N}\sum_{i=1}^{N}\delta(\boldsymbol{x}-\boldsymbol{x}^{(i)}) \tag{15-5}$$

可以将其看作离散采样的密度，因为

$$\int_{\mathbf{R}^{n_x}} p_s(\boldsymbol{x})\mathrm{d}\boldsymbol{x} = \frac{1}{N}\sum_{i=1}^{N}\int_{\mathbf{R}^{n_x}}\delta(\boldsymbol{x}-\boldsymbol{x}^{(i)})\mathrm{d}\boldsymbol{x} = 1 \tag{15-6}$$

图 15-2 中的直方图是块状的，在每个条的边界处不连续。此外，它的导数只在条的边界处非零，在其他位置处处为零。这个困难可以通过对 $p_s(\boldsymbol{x})$ 用一个平滑多尺度核函数$\mathcal{K}_h(x)$进行卷积，以代替直方图估计器来解决，即

$$\begin{aligned}\hat{p}(\boldsymbol{x}) &= (p_s * \mathcal{K}_h)(\boldsymbol{x}) \triangleq \int_{\mathbf{R}^{n_x}} p_s(\boldsymbol{u})\, \mathcal{K}_{\boldsymbol{H}}(\boldsymbol{x}-\boldsymbol{u})\mathrm{d}\boldsymbol{u} \\ &= \int_{\mathbf{R}^{n_x}} \frac{1}{N}\sum_{i=1}^{N} \delta(\boldsymbol{u}-\boldsymbol{x}^{(i)})\, \mathcal{K}_{\boldsymbol{H}}(\boldsymbol{x}-\boldsymbol{u})\mathrm{d}\boldsymbol{u} = \frac{1}{N}\sum_{i=1}^{N} \mathcal{K}_{\boldsymbol{H}}(\boldsymbol{x}-\boldsymbol{x}^{(i)})\end{aligned} \tag{15-7}$$

根据核的定义，多尺度核定义如下

$$\mathcal{K}_{\boldsymbol{H}}(\boldsymbol{t}) \triangleq \frac{1}{|\boldsymbol{H}|^{1/2}} \mathcal{K}(\boldsymbol{H}^{-1/2}\boldsymbol{t}) \tag{15-8}$$

这里$\mathcal{K}(\boldsymbol{u})$为一个连续对称核函数，满足条件

$$\int_{\mathbf{R}^n} \mathcal{K}(\boldsymbol{u})\mathrm{d}\boldsymbol{u} = 1 \tag{15-9}$$

$$\int_{\mathbf{R}^n} u^j\, \mathcal{K}(\boldsymbol{u})\mathrm{d}\boldsymbol{u} = 0 \quad j = 1,\cdots,k-1 \tag{15-10}$$

$$\int_{\mathbf{R}^n} u^k\, \mathcal{K}(\boldsymbol{u})\mathrm{d}\boldsymbol{u} \neq 0 \tag{15-11}$$

$$\mathcal{K}(\boldsymbol{u}) > 0, \forall u \tag{15-12}$$

式中：k 为核函数的阶数。矩阵 $\boldsymbol{H}$ 的元素为核函数的带宽（或平滑）参数，$\boldsymbol{H}$ 函数形式与协方差矩阵相似。Fukunaga[51] 和 Silverman[6] 的专题论文是非常好的介绍核密度估计的文献。

15.1.3.1 单变量核密度估计

首先考虑单变量情况，文献[7]中研究了多种单变量核函数，表 15-1 中列举了其中最为常用的。表 15-1 中指标函数 I 的定义如下

$$I(|u| \leqslant 1) = \begin{cases} 1 & -1 \leqslant u \leqslant 1 \\ 0 & \text{其他} \end{cases} \tag{15-13}$$

为表明核函数在一维密度估计中作用，考虑 12 个一维数据点的小集合 $\{-7.5, -6, -3, -2, -1.5, -1, -0.3, 0, 3.5, 4, 6, 8\}$。设标记的原始数据点的数量为 N_{old}，将表 15-1 的高斯核代入式(15-7)，得核密度估计为

$$\begin{aligned}\hat{p}(x) &= \frac{1}{N}\sum_{i=1}^{N} \mathcal{K}_h(x-x^{(i)}) = \sum_{i=1}^{N} \frac{1}{Nh}\mathcal{K}\left(\frac{x-x^{(i)}}{h}\right) \\ &= \sum_{i=1}^{N} \frac{1}{N}\frac{1}{\sqrt{2\pi}h}\exp\left(-\frac{1}{2}\left[\frac{x-x^{(i)}}{h}\right]^2\right)\end{aligned} \tag{15-14}$$

表 15－1　常用 2 阶单变量核函数

核函数	$\mathcal{K}(u)$
Uni form	$\frac{1}{2}I(\|u\|\leqslant 1)$
Triangle	$(1-\|u\|)I(\|u\|\leqslant 1)$
Epanechnikov	$\frac{3}{4}(1-u^2)I(\|u\|\leqslant 1)$
Quartic	$\frac{15}{16}(1-u^2)^2I(\|u\|\leqslant 1)$
Triweight	$\frac{35}{32}(1-u^2)^3I(\|u\|\leqslant 1)$
Cosine	$\frac{\pi}{4}\cos\left(\frac{\pi}{2}u\right)I(\|u\|\leqslant 1)$
Gaussian	$\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}u^2\right)$

很容易看出，带宽参数 h 正好是高斯密度的标准偏差。还可以看到，密度估计 $\hat{p}(x)$ 只不过是一个高斯混合密度，对于每个单独以初始采样点为中心高斯密度取相同的权值。

将单个采样点核函数定义为

$$K_i(x)=\frac{1}{N}\frac{1}{\sqrt{2\pi}h}\exp\left(-\frac{1}{2}\left[\frac{x-x^{(i)}}{h}\right]^2\right)=\frac{1}{N}\mathcal{N}(x;x^{(i)},h^2) \tag{15-15}$$

图 15－3 绘制了两个不同 h 值的完整估计核密度和单个的核函数，其中 x 的取值范围 $-15\leqslant x\leqslant 15$。

由图 15－3 可见，h 值的选取非常重要。在第一个图中，取 $h=0.9$，密度是平滑的，并且很好地逼近了采样点密度；第二个图中，取 $h=0.3$，密度估计几乎在每个采样点都有一个峰值，很不平滑。有评论认为 h 值的选取比 $\mathcal{K}$ 值选取重要得多[7]。

假设原始样本集来自于高斯分布（当然实际情况未必如此），即假设真实密度可以由 $\mathcal{N}(x;x_s,\sigma_s^2)$ 近似，这里 x_s 和 σ_s^2 分别为样本均值和方差。如果进行如下变换

$$\tilde{x}=\frac{x}{\sigma_s} \tag{15-16}$$

得

$$\mathcal{N}(x;x_s,\sigma_s^2)=\frac{1}{\sigma_s}\mathcal{N}(\tilde{x};\tilde{x}_s,1)=\frac{1}{\sigma_s}\mathcal{N}\left(\frac{x}{\sigma_s};\frac{x_s}{\sigma_s},1\right) \tag{15-17}$$

这样式（15－15）变成

$$K_i(\tilde{x})=\frac{1}{N\sigma_s}\mathcal{N}(x;x^{(i)},\sigma_s^2h^2)=\frac{1}{N}\frac{1}{\sqrt{2\pi}h\sigma_s}\exp\left(-\frac{1}{2}\left[\frac{x-x^{(i)}}{\sigma_sh}\right]^2\right) \tag{15-18}$$

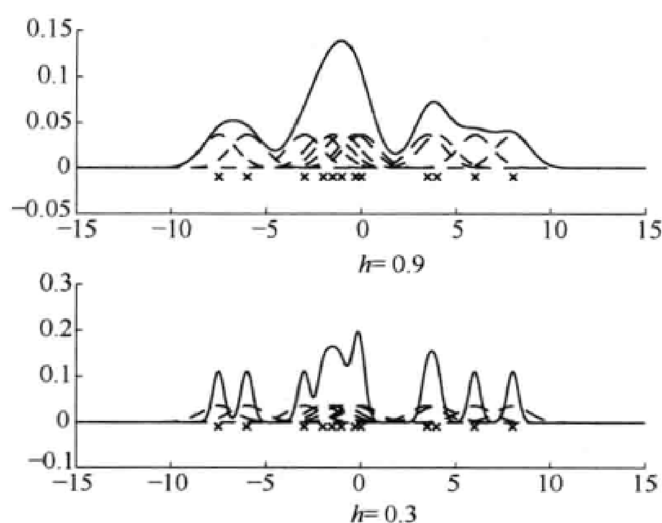


图 15－3　稀疏采样数据的高斯核密度估计

进行仿射变换

$$\tilde{x} = x^{(i)} + \sigma_s hc \tag{15-19}$$

对式(15－18)进行修正,得

$$K_i(\tilde{x}) = \frac{1}{N}\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}c^2\right) \tag{15-20}$$

这样式(15－14)简化为

$$\hat{p}(\tilde{x}) = \sum_{i=1}^{N}\frac{1}{N}\mathcal{N}(c;0,1) \tag{15-21}$$

这样可得一组新的样本限$\{x^{(j)}, j=1,\cdots,N_{new}\}$,这里$N_{new}$可以远远大于$N_{old}$,可以根据表 15－2 中的步骤得$\hat{p}(\tilde{x})$值。

表 15－2　生成一维核函数采样点

(1)将区间(0,1]细分为N_{old}个增量	$q_i=(i-1)/N_{old}\quad i=1,\cdots,N_{old}+1$
(2)由一致性分布生成N_{new}个样本	$u_j\sim u(0,1]\quad j=1,\cdots,N_{new}$
(3)对于u_j,决定其合适的q_i区间	找到满足$q_i<u_j\leqslant q_{i+1}$的i值
(4)生成c的样本	$c^{(j)}\sim\mathcal{N}(c;0,1)$
(5)生成x的新样本	$\tilde{x}^{(j)}=x^{(i)}+\sigma_s hc^{(j)}$
(6)对于$j=1,\cdots,N_{new}$,重复步骤(2)～(5)	

因为这里利用了核估计方法与混合密度方法得的权重相同这一事实,正则化过程变得非常简单。下面将给出推广到包括不相等权值的多变量情况。注

意，高密度采样区域会选到更多的重采样值。

应该指出的是，采用高斯分布之外的核函数时，核密度估计的重采样有时困难得多。例如，由 Epanechnikov 核函数生成样本包括生成$\sqrt{\beta}u$ 步骤，这里 β 服从参数为(1/2,2)的 β 分布、u 为均匀分布的采样，参见上面步骤(2)[8]。

关于多维核方法的进一步讨论可以参阅文献[9－12]。

15.1.3.2 多变量核密度估计

对于多参数高斯核函数，由式(15－18)可得密度估计器为

$$\hat{p}(\boldsymbol{x}) = \sum_{i=1}^{N} \frac{1}{N} \frac{1}{(2\pi)^{n_x/2} |\Sigma \boldsymbol{H}|^{1/2}} \exp\left\{ -\frac{1}{2}(\boldsymbol{x} - \boldsymbol{x}^{(i)})^{\mathrm{T}} [\Sigma \boldsymbol{H}]^{-1} (\boldsymbol{x} - \boldsymbol{x}^{(i)}) \right\} \tag{15-22}$$

很显然，多参数单个核函数为高斯密度$\mathcal{N}(\boldsymbol{x};\boldsymbol{x}^{(i)},\Sigma\boldsymbol{H})$除以 N。

考虑 $n_x=2$ 的二维情况，其中

$$\boldsymbol{H} = \begin{bmatrix} h_1^2 & h_{12} \\ h_{12} & h_2^2 \end{bmatrix} \tag{15-23}$$

在这种情况下，可以有 3 个不同的平滑参数：一个用于维度 1；一个用于维度 2；还有一个用于两个维度之间的相关性。但与一维情况相比，选择“最优”的平滑参数更为复杂。假设数据密度为高斯分布，即 $\boldsymbol{x} \sim \mathcal{N}(\boldsymbol{x};\hat{\boldsymbol{x}},\Sigma)$，可以由该数据生成一个样本协方差矩阵 Σ_s，对其进行矩阵分解 $\Sigma_s = \boldsymbol{D}_s\boldsymbol{D}_s^{\mathrm{T}}$，可利用如下变换实现数据规范化，即

$$\tilde{\boldsymbol{x}} = \boldsymbol{D}_s^{-1}\boldsymbol{x} \tag{15-24}$$

得

$$\tilde{\boldsymbol{x}} \sim |\boldsymbol{D}_s|^{-1} \mathcal{N}(\boldsymbol{D}_s^{-1}\boldsymbol{x};\boldsymbol{D}_s^{-1}\boldsymbol{x}_s,\boldsymbol{I}) \tag{15-25}$$

由于变换后的数据有一个单位对角采样协方差矩阵，可将 $\boldsymbol{H}$ 的非对角线项设为零，并且设 $h=h_1=h_2$，因此有 $|\boldsymbol{H}|^{1/2}=h^2$。这样，对于二维情况，式(15－22)变成

$$\hat{p}(\tilde{\boldsymbol{x}}) = \sum_{i=1}^{N} \frac{1}{N} \frac{1}{2\pi |\boldsymbol{D}_s| h^2} \exp\left\{ -\frac{1}{2}(\boldsymbol{D}_s^{-1}\boldsymbol{x} - \boldsymbol{D}_s^{-1}\boldsymbol{x}^{(i)})^{\mathrm{T}} \boldsymbol{H}^{-1} \times (\boldsymbol{D}_s^{-1}\boldsymbol{x} - \boldsymbol{D}_s^{-1}\boldsymbol{x}^{(i)}) \right\} \tag{15-26}$$

如果最后实施的仿射变换为

$$\boldsymbol{x} = \boldsymbol{x}^{(i)} + \boldsymbol{D}_s h \boldsymbol{c} \tag{15-27}$$

那么式(15－26)简化为

$$\hat{p}(\tilde{x}) = \sum_{i=1}^{N} \frac{1}{N} \mathcal{N}(\boldsymbol{c};\boldsymbol{0},\boldsymbol{I}) \tag{15-28}$$

将以上推导推广至2维以上非常简单，这里作为练习留给读者。

可以从MATLAB中心[12]下载二维高斯核密度估计子程序，其中算法经过改进后，可实现由图15－1中的数据生成平滑的密度估计。图15－4给出估计pdf的可视化结果。

与一维情况类似，按照表15－3中的步骤，很容易得到多维高斯核函数采样。

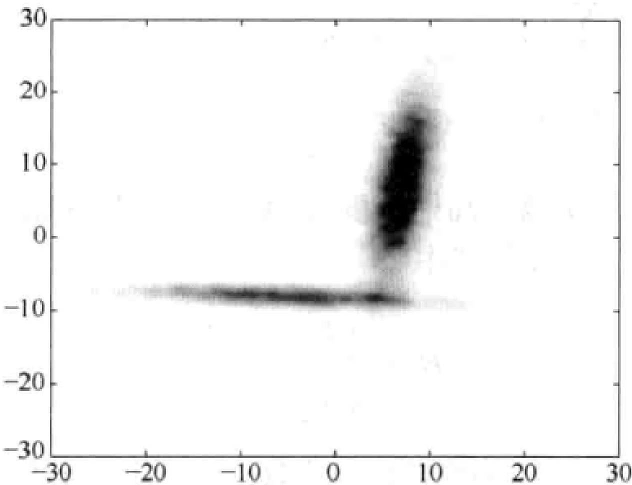


图15－4　由密度函数随机采样生成的高斯核密度估计的可视化

表15－3　生成多维核采样

(1)将区间(0,1]细分为N_{old}个增量	$q_i=(i-1)/N_{old}\quad i=1,\cdots,N_{old}+1$
(2)由一致性分布生成一个样本	$u_j \sim u(0,1]\quad j=1,\cdots,N_{new}$
(3)对于u_j，决定其合适的q_i区间	找到满足$q_i<u_j\leqslant q_{i+1}$的i值
(4)生成c的样本	Select $\boldsymbol{x}^{(i)}$
(5)生成x的新样本	$\boldsymbol{c}^{(j)} \sim \mathcal{N}(\boldsymbol{c};\boldsymbol{0},\boldsymbol{I})$
(6)对于$j=1,\cdots,N_{new}$，重复步骤(2)～(5)	$\boldsymbol{x}^{(j)}=\boldsymbol{x}^{(i)}+\boldsymbol{D}_s hc^{(j)}$

即，由一组随机样本$\{\boldsymbol{x}^{(i)},i=1,\cdots,N_{old}\}$，可产生一个新的样本集合$\{\boldsymbol{x}^{(j)},j=1,\cdots,N_{new}\}$，这里$N_{new}$可以远远大于$N_{old}$。

15.2　重要性采样的基本概念

由本书第Ⅱ部分可知，几乎所有的非线性系统高斯估计(跟踪)方法都涉及某种密度加权矩积分的数值计算。其中一些数值方法包括函数在特定确定性Sigma点取值的加权和。

如果加权密度是多参数、非高斯或未知的，除个别分析密度已知的情况[13]，一般不能使用确定性数值方法。蒙特卡洛方法在一般密度加权积分方面最近（近 30 年来）的研究进展，已经广泛地应用于估计和跟踪当中。这些方法中最具影响力的是独立重要性抽样。

读者可以参考 Robert、Casella[14] 和 Doucet 等[15] 的书籍，还有 Doucet[16] 和 Liu、Chen[17] 的论文，及其包含的参考文献。其中有一般蒙特卡洛积分方法和序贯重要性采样方法的起源和基本发展的完整阐述。

第 12 章用高斯密度生成蒙特卡洛采样来求解高斯密度加权积分，将积分简化为函数在采样点取值的加权和。对于未知密度或非高斯密度，由该密度生成蒙特卡洛样本过程可以用重要性采样来高效实现。在重要性采样中，由已知重要性密度 $q(\boldsymbol{x})$ 生成一组独立的蒙特卡洛采样 $\{\boldsymbol{x}^{(1)},\cdots,\boldsymbol{x}^{(N_s)}\}$，然后进行被积函数采样点取值加权和（平均）计算。$q(\boldsymbol{x})$ 的取值自由度很大，最好选择所生成的采样值落在对于积分值很重要区域，不在这个区域的被积函数是可以忽略不计的。

重要性采样的概念[13]源自于如下形式的积分，即

$$\hat{\boldsymbol{f}}(\boldsymbol{x}) = \int \boldsymbol{f}(\boldsymbol{x}) p(\boldsymbol{x}) \mathrm{d}\boldsymbol{x} \tag{15-29}$$

首先考虑 $p(\boldsymbol{x})$ 为高斯分布的情况。一种逼近积分的方法是由高斯分布 $\mathcal{N}(\boldsymbol{x};\boldsymbol{x},\boldsymbol{P}_x)$ 生成一组采样值 $\{\boldsymbol{x}^{(i)},i=1,\cdots,N\}$，用 $p(\boldsymbol{x})$ 的离散采样代替函数本身，即

$$p(\boldsymbol{x}) \simeq \frac{1}{N}\sum_{i=1}^{N}\delta(\boldsymbol{x}-\boldsymbol{x}^{(i)}) = \sum_{i=1}^{N} w_i \delta(\boldsymbol{x}-\boldsymbol{x}^{(i)}) \tag{15-30}$$

这里 $\omega_i = 1/N, \forall i$。式（15－29）的积分变成

$$\hat{\boldsymbol{f}}(\boldsymbol{x}) \simeq \int \boldsymbol{f}(\boldsymbol{x}) \sum_{i=1}^{N} w_i \delta(\boldsymbol{x}-\boldsymbol{x}^{(i)}) \mathrm{d}\boldsymbol{x} = \sum_{i=1}^{N} w_i \boldsymbol{f}(\boldsymbol{x}^{(i)}) \tag{15-31}$$

式中：ω_i 表示 $\boldsymbol{x}^{(i)}$ 相对于样本均值（总和）的重要性。因为可以直接从已知密度中进行采样，在高斯密度情况下，所有样本相对于样本均值总和的重要性相同，因此 $\omega_i = 1/N$。这里权值 ω_i 不代表 $\boldsymbol{x}^{(i)}$ 的概率，只是表示用加权和逼近矩积分时 $\boldsymbol{x}^{(i)}$ 的重要性。文献[13]中指出，当 $N\to\infty$ 时，式（15－31）收敛于精确解。

假设 $p(\boldsymbol{x})$ 的 PDF 未知或采样困难，再假设有另一个已知或容易采样的 PDF$q(x)$，称之为重要性密度，后面将解释这样命名的原因。举例说明，$p(\boldsymbol{x})$ 为未知密度，$q(x)$ 为高斯密度。假设 $q(x)$ 的支撑大于或等于 $p(x)$ 的支撑，这样在每个 x 值处，$p(\boldsymbol{x})$ 与 $q(\boldsymbol{x})$ 成比例。由于 $p(\boldsymbol{x})$ 是一个规范化的 PDF，$q(x)$ 必须是 $p(\boldsymbol{x})$ 的比例缩放，对于每个 $\boldsymbol{x}$，有不同的尺度因子。对于任意两个具有相同支撑的 PDF，这是绝对正确的。现在，可以将比例或权重因子定义为

$$w(\boldsymbol{x}) = \frac{p(\boldsymbol{x})}{q(\boldsymbol{x})} \tag{15-32}$$

必要条件是 $q(\boldsymbol{x})$ 的支撑大于或等于 $p(\boldsymbol{x})$ 的支撑。因此，在许多情况下，当 $\boldsymbol{x}$ 的取值为 $-\infty \leqslant x_i \leqslant \infty, i=1,\cdots,n_x$ 时，$q(\boldsymbol{x})$ 最好的选择是多维高斯分布。

图 15－5 给出了一维的例子，其中 $\hat{p}(x)$（上面图中的实线）为多峰的，$q(x)$（上面图中的虚线）为比 $p(x)$ 更宽的高斯 PDF，在图（15－5）中给出了 $\omega(x)$ 生成示例。

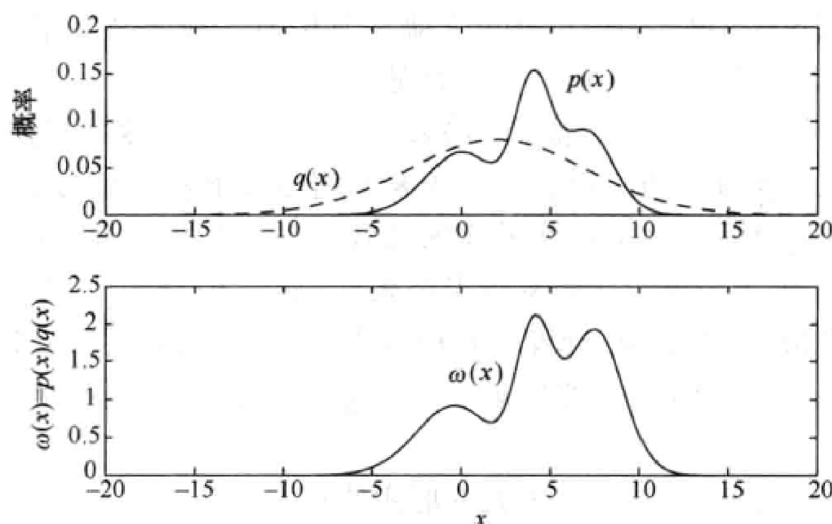


图 15－5　$\omega(x)$ 生成示例

必须强调，$\omega(\boldsymbol{x})$ 不是密度函数，而是两个密度的非规范化比值。即使在规范化后，$\omega(\boldsymbol{x})$ 也不能取代 $p(\boldsymbol{x})$。在使用 $\boldsymbol{x}$ 的最大后验估计（MAP）时，这一点很重要。如果 $q(\boldsymbol{x})$ 的最大值与 $p(\boldsymbol{x})$ 的最大值不一致，那么 $\omega(\boldsymbol{x})$ 的最大值就不是 $p(\boldsymbol{x})$ 的最大后验估计。

假设 $q(x)$ 为高斯分布，有 $q(x)=\mathcal{N}(x;\mu,\sigma^2)$。根据 $q(x)$ 的采样值，$x^{(i)}$ 可以求得 x_i 的概率 $p(x)$，即

$$q(x=x^{(i)}) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left\{-\frac{1}{2\sigma^2}((x=x^{(i)})-\mu)^2\right\} \tag{15-33}$$

从而有

$$w(x=x^{(i)}) = \frac{p(x=x^{(i)})}{q(x=x^{(i)})} \tag{15-34}$$

对于已知分布 $q(\boldsymbol{x})$ 的多维采样 $\{\boldsymbol{x}^{(i)}, i=1,\cdots,N\}$，采用某种有待确定的方法来为每个样本生成 ω_i，用 $p(\boldsymbol{x})$ 的离散值取代函数本身，即

$$p(\boldsymbol{x}) \simeq \sum_{i=1}^{N} w_i \delta(\boldsymbol{x}-\boldsymbol{x}^{(i)}) \tag{15-35}$$

这样积分可以被重写为

$$\hat{\boldsymbol{f}}(\boldsymbol{x}) = \int \boldsymbol{f}(\boldsymbol{x})p(\boldsymbol{x})\mathrm{d}\boldsymbol{x} = \int \boldsymbol{f}(\boldsymbol{x})w(\boldsymbol{x})q(\boldsymbol{x})\mathrm{d}\boldsymbol{x}$$

$$= \int \boldsymbol{f}(\boldsymbol{x}) \sum_{i=1}^{N} w_i \delta(\boldsymbol{x} - \boldsymbol{x}^{(i)}) \mathrm{d}\boldsymbol{x} = \sum_{i=1}^{N} w_i \boldsymbol{f}(\boldsymbol{x}^{(i)}) \quad (15-36)$$

上式与式(15－31)的形式完全相同,但权重 ω_i 的取值不再统一等于 $1/N$。由于权重是式(15－35)离散密度逼近定义的一部分,很显然必须对权重进行规范化,使得 $\Sigma\omega_i = 1$。

考虑一阶马尔可夫过程 $\boldsymbol{x}_n$,其估计矩由条件后验密度 $p(\boldsymbol{x}_n | \boldsymbol{z}_{1:n})$ 确定,其中 $p(\boldsymbol{x}_0)$ 为得任何观测值之前的初始后验概率(参见 3.3)。给定任意后验 PDF,矩的非线性函数通用表示形式为

$$\boldsymbol{I}(\boldsymbol{x}_n) = \int \boldsymbol{g}(\boldsymbol{x}_n) p(\boldsymbol{x}_n \mid z_{1:n}) \mathrm{d}\boldsymbol{x}_n \quad (15-37)$$

这里 $\boldsymbol{g}(\boldsymbol{x}_n)$ 是一个函数,可以表示任意矩生成函数。例如,如果要求得函数 $\boldsymbol{f}(\boldsymbol{x}_n)$ 的均值,那么由 $\boldsymbol{g}(\boldsymbol{x}_n) = \boldsymbol{f}(\boldsymbol{x}_n)$ 可导出 $\boldsymbol{I}(\boldsymbol{x}_n) = \hat{\boldsymbol{f}}(\boldsymbol{x}_n)$。如果要计算 $\boldsymbol{f}(\boldsymbol{x}_n)$ 的协方差,令 $\boldsymbol{g}(\boldsymbol{x}_n) = [\boldsymbol{f}(\boldsymbol{x}_n) - \hat{\boldsymbol{f}}(\boldsymbol{x}_n)][\boldsymbol{f}(\boldsymbol{x}_n) - \hat{\boldsymbol{f}}(\boldsymbol{x}_n)]^{\mathrm{T}}$,可得 $\boldsymbol{I}(\boldsymbol{x}_n) = \boldsymbol{P}_n^{ff}$。

假设 $p(\boldsymbol{x}_n | z_{1:n})$ 未知或采样困难,但是有一个很容易的采样重要性密度 $q(\boldsymbol{x}_n | z_{1:n})$,与 $p(\boldsymbol{x}_n | z_{1:n})$ 的支撑相同,其最大值与 $p(\boldsymbol{x}_n | z_{1:n})$ 在大致相同的位置,式(15－37)可写为

$$\begin{aligned} \boldsymbol{I}(\boldsymbol{x}_n) &= \int \boldsymbol{g}(\boldsymbol{x}_n) \frac{p(\boldsymbol{x}_n \mid z_{1:n})}{q(\boldsymbol{x}_n \mid z_{1:n})} q(\boldsymbol{x}_n \mid z_{1:n}) \mathrm{d}\boldsymbol{x}_n \\ &= \int \boldsymbol{g}(\boldsymbol{x}_n) w(\boldsymbol{x}_n) q(\boldsymbol{x}_n \mid z_{1:n}) \mathrm{d}\boldsymbol{x}_n \end{aligned} \quad (15-38)$$

这里权重函数 $\omega(\boldsymbol{x}_n)$ 定义为

$$w(\boldsymbol{x}_n) \triangleq \frac{p(\boldsymbol{x}_n \mid z_{1:n})}{q(\boldsymbol{x}_n \mid z_{1:n})} \quad (15-39)$$

对 $p(\boldsymbol{x}_n | z_{1:n})$ 使用式(3－23)贝叶斯法则,式(15－39)变为

$$w(\boldsymbol{x}_n) = \frac{p(z_n \mid \boldsymbol{x}_n) p(\boldsymbol{x}_n \mid z_{1:n-1})}{p(z_n \mid z_{1:n-1})} \frac{1}{q(\boldsymbol{x}_n \mid z_{1:n})} \quad (15-40)$$

注意 $p(z_n | z_{1:n-1})$ 恰好为正则项,得

$$w(\boldsymbol{x}_n) \propto \frac{p(z_n \mid \boldsymbol{x}_n) p(\boldsymbol{x}_n \mid z_{1:n-1})}{q(\boldsymbol{x}_n \mid z_{1:n})} \quad (15-41)$$

这里 $\propto$ 表示成比例。根据似然函数,$p(z_n | \boldsymbol{x}_n)$、先验 PDF$p(\boldsymbol{x}_n | z_{1:n-1})$ 和重要性 PDF$q(\boldsymbol{x}_n | z_{1:n})$,式(15－41)给出一般加权函数。

如果选择 $\{\boldsymbol{x}_n^{(i)}, i = 1, \cdots, N_s\}$ 样本集合,可通过狄拉克 δ 函数组(见式(15－5)),得 $q(\boldsymbol{x}_n | z_{1:n})$ 估计为

$$q(\boldsymbol{x}_n \mid z_{1:n}) = \frac{1}{N_s} \sum_{i=1}^{N_s} \delta(\boldsymbol{x}_n - \boldsymbol{x}_n^{(i)}) \quad (15-42)$$

式(15－38)变成

$$I(\boldsymbol{x}_n) = \int \boldsymbol{g}(\boldsymbol{x}_n) w(\boldsymbol{x}_n) \frac{1}{N_s}\sum_{i=1}^{N_s}\delta(\boldsymbol{x}_n - \boldsymbol{x}_n^{(i)})\mathrm{d}\boldsymbol{x}_n = \frac{1}{N_s}\sum_{i=1}^{N_s} w_n^{(i)}\boldsymbol{g}(\boldsymbol{x}_n^{(i)}) \tag{15-43}$$

$$\omega_n^{(i)} \triangleq \omega(\boldsymbol{x}_n^{(i)})$$

由式(15－39),得

$$w_n^{(i)} = \frac{p(\boldsymbol{x}_n^{(i)} \mid \boldsymbol{z}_{1:n})}{q(\boldsymbol{x}_n^{(i)} \mid \boldsymbol{z}_{1:n})} \tag{15-44}$$

或者,利用式(15－41),得一组非规范化权值,即

$$\widetilde{w}_n^{(i)} = \frac{p(\boldsymbol{z}_n \mid \boldsymbol{x}_n^{(i)})p(\boldsymbol{x}_n^{(i)} \mid \boldsymbol{z}_{1:n-1})}{q(\boldsymbol{x}_n^{(i)} \mid \boldsymbol{z}_{1:n})} \tag{15-45}$$

由于将式(15－41)中的成比例替换为相等,因此权值是非规范化的,可以通过令其总和为1来归一化权重,即

$$w_n^{(i)} \triangleq \frac{1/N_s\, \widetilde{w}_n^{(i)}}{1/N_s \sum_{i=1}^{N_s} \widetilde{w}_n^{(i)}} \tag{15-46}$$

这样式(15－43)变成

$$I(\boldsymbol{x}_n) = \sum_{i=1}^{N_s} w_n^{(i)}\boldsymbol{g}(\boldsymbol{x}_n^{(i)}) \tag{15-47}$$

这里再着重强调一次,$\omega_n^{(i)}$ 不是 $\boldsymbol{x}_n^{(i)}$ 的概率,而是 $\boldsymbol{x}_n^{(i)}$ 在离散矩积分和中的重要度。

15.3 总结

本章提出了两个主要概念:一个是基于一组密度函数采样集,来表示该密度函数的方法;另一个是函数数值积分中,由任意 PDF 给定权重的函数的蒙特卡洛重要性采样的概念。目前,仅仅基于这些结果,还未能提出一种估计和跟踪的实用方法,还需要进行进一步的改进。

基于顺序生成的蒙特卡洛点和权重的递归集合$\{\boldsymbol{x}_n^{(i)}, \omega_n^{(i)}, i=1,\cdots,N_s, n=1,\cdots,N\}$,所得的跟踪滤波器被称为序贯蒙特卡洛(Sequential Monte Carlo,SMC)方法。与现有的其他技术相比,SMC 方法具有诸多优点:①它允许在一般状态空间模型,包括非线性非高斯条件下的后验分布推理,因此利用 SMC 方法可以计算各种矩、分位数和最大后验密度区域,而高斯卡尔曼滤波器仅能进行一阶、

二阶矩的逼近;②通过指定合适的状态空间模型,SMC 方法可结合状态空间约束,如用限制状态空间的范围来表示物理限制,例如速度约束或硬边界。

下一章将介绍一个与 SMC 相比约束更大的方法,用顺序的方式为每个采样点生成一个权重,得序贯重要性采样(SIS)粒子滤波器。第 17 章将介绍一种约束较小的方法,但是在每次迭代中,需要重新计算权重函数。

参考文献

[1] Richardson HR, Stone LD, Monach WR, et al. Early maritime applications of particle filtering. *Signal and Data Processing of Small Targets*, *Proceedings of SPIE*, 2003,5204: 165 - 174.

[2] Agnew DC, Constable C. *Geophysical Data Analy sis*. [2005]. [Online] http://mahi. ucsd. edu/cathy/Classes/SI0223/Partl/. Chapter 9.

[3] Scott DW. Multivariate density estimation and visualization. In *Handbook of Statistics*, *Vol.* 23: *Data Mining and Computational Statistics*. Springer,2004.

[4] Patlolla R. 2 - Dimensional histogram. [Online] http://www. mathworks. com/matlabcentral/ fileexchange/8422.

[5] Fukunaga K. Introduction to Statistical Pattern Recognition. 2nd ed. Academic Press, 1990.

[6] Silverman BW. Density Estimation for Statistics and Data Analysis. New York: Chapman and Hall,1986.

[7] Turlach BA. Bandwidth selection in Kernel density estimation: a review. *In CORE Institut Stat*,1993:23 - 49.

[8] Musso C, Oudjane N, Le Gland F. Improving regularized particle filters. In Doucet A, de Freita N, Gordon G, editors. *Sequential Monte Carlo Methods in Practice*. Springer,2001.

[9] Wand MP, Jones MC. Comparison of smoothing parameterization in bivariate Kernel density estimation. *J. Am. Stat. Soc*,1993,88(422):520 - 528.

[10] Zhang X, King ML, Hyndman RJ. Bandwidth selection for multivariate Kernel density estimation using MCMC. *Comp. Stat. Data Anal*, 2006,50:3009 - 3031.

[11] Wikipedia. [Online] http://en. wikipedia. org/wiki/Kernel. (statistics).

[12] Botev Z. Two - Dimensional Kernel Density Estimator Function. [Online] http://www. mathworks. com/matlabcentral/fileexchange/17204 - kernel - density - estimation.

[13] Evans M, Swartz T. Approximating Integrals via Monte Carlo and Deterministic Methods. New York: Oxford University Press,2005.

[14] Robert CP, Casella G. *Monte Carlo Statistical Methods*. Springer,2004.

[15] Doucet A, de Freita N, Gordon G. *Sequential Monte Carlo Methods in Practice*. Springer,2001.

[16] Doucet A. On Sequential Simulation - Based Methods for Bayesian Estimation. Technical Report CUED/F - INFENGATR. 310,1998.

[17] Liu JS, Chen R. Sequential Monte Carlo methods for dynamic systems. *J. Am. Stat. Soc*,1998, 93(443): 1032 - 1044.

第16章 序贯重要性采样粒子滤波器

16.1 序贯重要性采样的基本概念

在第Ⅱ部分提出的所有高斯卡尔曼滤波估计器中，t_n 时刻的估计值以一种解析方式依赖于 t_{n-1}时刻的估计值。本章要找到一个类似的递归公式，当噪声密度为非高斯时，用重要性采样方法进行估计。

基于权值序列估计的递归估计方法被称为序贯重要性采样(SIS)粒子滤波器。利用式(3－24)的 Chapman－Kolmogorov 定理，可将式(15－41)中的 $p(\boldsymbol{x}_n|\boldsymbol{z}_{1:n-1})$和 $q(\boldsymbol{x}_n|\boldsymbol{z}_{1:n})$进一步展开，这样式(15－41)变成

$$w(\boldsymbol{x}_n) \propto \frac{p(\boldsymbol{z}_n \mid \boldsymbol{x}_n)\int p(\boldsymbol{x}_n \mid \boldsymbol{x}_{n-1},\boldsymbol{z}_{1:n-1})p(\boldsymbol{x}_{n-1} \mid \boldsymbol{z}_{1:n-1})\mathrm{d}\boldsymbol{x}_{n-1}}{\int q(\boldsymbol{x}_n \mid \boldsymbol{x}_{n-1},\boldsymbol{z}_{1:n})q(\boldsymbol{x}_{n-1} \mid \boldsymbol{z}_{1:n-1})\mathrm{d}\boldsymbol{x}_{n-1}} \tag{16-1}$$

在分母中包含 $q(\boldsymbol{x}_{n-1}|\boldsymbol{z}_{1:n-1})$而不是 $q(\boldsymbol{x}_{n-1}|\boldsymbol{z}_{1:n})$的原因在于：$\boldsymbol{x}_{n-1}$不能以未来的观测作为条件。假定状态演化方程是独立于观测值的一阶马尔可夫过程，令 $p(\boldsymbol{x}_n|\boldsymbol{x}_{n-1},\boldsymbol{z}_{1:n-1})\rightarrow p(\boldsymbol{x}_n|\boldsymbol{x}_{n-1})$，得

$$w(\boldsymbol{x}_n) \propto \frac{p(\boldsymbol{z}_n \mid \boldsymbol{x}_n)\int p(\boldsymbol{x}_n \mid \boldsymbol{x}_{n-1})p(\boldsymbol{x}_{n-1} \mid \boldsymbol{z}_{1:n-1})\mathrm{d}\boldsymbol{x}_{n-1}}{\int q(\boldsymbol{x}_n \mid \boldsymbol{x}_{n-1},\boldsymbol{z}_{1:n})q(\boldsymbol{x}_{n-1} \mid \boldsymbol{z}_{1:n-1})\mathrm{d}\boldsymbol{x}_{n-1}} \tag{16-2}$$

对于马尔可夫过程，可以进行如下假设，即

$$q(\boldsymbol{x}_n \mid \boldsymbol{x}_{n-1},\boldsymbol{z}_{1:n}) = q(\boldsymbol{x}_n \mid \boldsymbol{x}_{n-1},\boldsymbol{z}_n) \tag{16-3}$$

这样，重要性密度仅取决于 $\boldsymbol{x}_{n-1}$和 $\boldsymbol{z}_n$。

假设有一组由 $q(\boldsymbol{x}_{n-1}|\boldsymbol{z}_{1:n-1})$得到(第15章由 $q(\boldsymbol{x}_n|\boldsymbol{z}_{1:n})$生成样本)随机样本$\{\boldsymbol{x}_{n-1}^{(i)},i=1,\cdots,N_s\}$，即$\{\boldsymbol{x}_{n-1}^{(i)}\sim q(\boldsymbol{x}_{n-1}|\boldsymbol{z}_{1:n-1})\}$，这样式(16－2)变成

$$\widetilde{w}(\boldsymbol{x}_n) = \sum_{i=1}^{N_s} \frac{p(z_n \mid \boldsymbol{x}_n)p(\boldsymbol{x}_n \mid \boldsymbol{x}_{n-1}^{(i)})p(\boldsymbol{x}_{n-1}^{(i)} \mid \boldsymbol{z}_{1:n-1})}{q(\boldsymbol{x}_n \mid \boldsymbol{x}_{n-1}^{(i)}, z_n)q(\boldsymbol{x}_{n-1}^{(i)} \mid \boldsymbol{z}_{1:n-1})} \tag{16-4}$$

根据式(15－39)所给出的 $w(\boldsymbol{x}_n)$ 的定义，得

$$\widetilde{w}(\boldsymbol{x}_n) = \sum_{i=1}^{N_s} w_{n-1}^{(i)} \frac{p(z_n \mid \boldsymbol{x}_n)p(\boldsymbol{x}_n \mid \boldsymbol{x}_{n-1}^{(i)})}{q(\boldsymbol{x}_n \mid \boldsymbol{x}_{n-1}^{(i)}, z_n)} \tag{16-5}$$

得如下的后验分布表达形式，即

$$p(\boldsymbol{x}_n \mid \boldsymbol{z}_{1:n}) = \sum_{i=1}^{N_s} w_{n-1}^{(i)} \frac{p(\boldsymbol{x}_n \mid \boldsymbol{x}_{n-1}^{(i)})}{q(\boldsymbol{x}_n \mid \boldsymbol{x}_{n-1}^{(i)}, z_n)} p(z_n \mid \boldsymbol{x}_n)q(\boldsymbol{x}_n \mid \boldsymbol{z}_{1:n}) \tag{16-6}$$

由 $q(\boldsymbol{x}_n \mid \boldsymbol{x}_{n-1}^{(i)}, z_n)$ 生成样本 $\{\boldsymbol{x}_n^{(i)}, i=1,\cdots,N_s\}$，即

$$\boldsymbol{x}_n^{(i)} \sim q(\boldsymbol{x}_n \mid \boldsymbol{x}_{n-1}^{(i)}, z_n) \tag{16-7}$$

$$q(\boldsymbol{x}_n \mid \boldsymbol{z}_{1:n}) = \frac{1}{N_s}\delta(\boldsymbol{x}_n - \boldsymbol{x}_n^{(i)}) \tag{16-8}$$

得

$$p(\boldsymbol{x}_n \mid \boldsymbol{z}_{1:n}) = \sum_{i=1}^{N_s} \frac{\widetilde{w}_{n-1}^{(i)}}{N_s} \frac{p(\boldsymbol{x}_n^{(i)} \mid \boldsymbol{x}_{n-1}^{(i)})}{q(\boldsymbol{x}_n^{(i)} \mid \boldsymbol{x}_{n-1}^{(i)}, z_n)} p(z_n \mid \boldsymbol{x}_n^{(i)})\delta(\boldsymbol{x}_n - \boldsymbol{x}_n^{(i)}) \tag{16-9}$$

用下式对 $\widetilde{w}_{n-1}^{(i)}/N_s$ 进行规范化，即

$$w_{n-1}^{(i)} = \frac{\widetilde{w}_{n-1}^{(i)}/N_s}{\sum_{i=1}^{N_s} \widetilde{w}_{n-1}^{(i)}/N_s} \tag{16-10}$$

得

$$p(\boldsymbol{x}_n \mid \boldsymbol{z}_{1:n}) = \sum_{i=1}^{N_s} w_n^{(i)}\delta(\boldsymbol{x}_n - \boldsymbol{x}_n^{(i)}) \tag{16-11}$$

序贯重要性递归权值更新方程为

$$w_n^{(i)} \triangleq w_{n-1}^{(i)} \frac{p(\boldsymbol{x}_n^{(i)} \mid \boldsymbol{x}_{n-1}^{(i)})}{q(\boldsymbol{x}_n^{(i)} \mid \boldsymbol{x}_{n-1}^{(i)}, z_n)} p(z_n \mid \boldsymbol{x}_n^{(i)}) \tag{16-12}$$

回到式(15－37)给出的一般矩方程，$\boldsymbol{x}_n$ 关于 $p(\boldsymbol{x}_n \mid z_{1:n})$ 的前两阶矩可以由下式求得，即

$$\begin{aligned}\hat{\boldsymbol{x}}_n \triangleq \varepsilon\{\boldsymbol{x}_n \mid \boldsymbol{z}_{1:n}\} &= \int_{\mathbf{R}} \boldsymbol{x}_n p(\boldsymbol{x}_n \mid \boldsymbol{z}_{1:n})\mathrm{d}\boldsymbol{x}_n = \int_{\mathbf{R}} \boldsymbol{x}_n \sum_{i=1}^{N_s} w_n^{(i)}\delta(\boldsymbol{x}_n - \boldsymbol{x}_n^{(i)}) \\ &= \sum_{i=1}^{N_s} w_n^{(i)}\boldsymbol{x}_n^{(i)}\end{aligned} \tag{16-13}$$

类似地有

$$P_n^{xx} \triangleq \varepsilon\{(\boldsymbol{x}_n - \hat{\boldsymbol{x}}_n)(\boldsymbol{x}_n - \hat{\boldsymbol{x}}_n)^{\mathrm{T}} \mid z_{1:n}\} = \sum_{i=1}^{N_s} w_n^{(i)} (\boldsymbol{x}_n^{(i)} - \hat{\boldsymbol{x}}_n)(\boldsymbol{x}_n^{(i)} - \hat{\boldsymbol{x}}_n)^{\mathrm{T}} \tag{16-14}$$

序贯重要性采样粒子滤波器的步骤非常简单,如表 16-1 所列。

表 16-1　一般序贯重要性采样粒子滤波器

1. 滤波器初始化 (1) 状态矢量样本初始化; (2) 权值初始化	$\boldsymbol{x}_0^{(i)} \sim q(\boldsymbol{x}_0) \quad i = 1,\cdots,N_s$ $w_0^{(i)} = \frac{1}{N_s}$
2. 序贯重要性采样 (1) 抽取新样本; (2) 生成非规范重要性权值; (3) 权值的规范化; (4) 时间递增,返回 1	$\boldsymbol{x}_n^{(i)} \sim q(\boldsymbol{x}_n \mid \boldsymbol{x}_{n-1}^{(i)}, z_n), \forall i$ $\tilde{w}_n^{(i)} = w_{n-1}^{(i)} \cdot \frac{p\, z_n \mid \boldsymbol{x}_n^{(i)})p(\boldsymbol{x}_n^{(i)} \mid \boldsymbol{x}_{n-1}^{(i)})}{q(\boldsymbol{x}_n^{(i)} \mid \boldsymbol{x}_{n-1}^{(i)} z_n)}$ $w_n^{(i)} = \tilde{w}_n^{(i)} / \sum_{i=1}^{N} \tilde{w}_n^{(i)}$ $\boldsymbol{x}_n^{(i)} \to \boldsymbol{x}_{n-1}^{(i)}$ $w_n^{(i)} \to w_{n-1}^{(i)}$
3. 如需要,计算矩 (1) 一阶矩:状态矢量均值; (2) 二阶矩:协方差	$\hat{\boldsymbol{x}}_n = \sum_{i=1}^{N_s} w_n^{(i)} \boldsymbol{x}_n^{(i)}$ $P_n^{xx} = \sum_{i=1}^{N_s} w_n^{(i)} \cdot (\boldsymbol{x}_n^{(i)} - \hat{\boldsymbol{x}}_n)(\boldsymbol{x}_n^{(i)} - \hat{\boldsymbol{x}}_n)^{\mathrm{T}}$

在序贯重要性采样粒子滤波过程中,包含几个重要的实际问题。第一个问题出现在序贯重要性采样的第 2 步,需要似然函数 $p(z_n \mid \boldsymbol{x}_n)$、预测密度 $p(\boldsymbol{x}_n \mid \boldsymbol{x}_{n-1})$ 和重要性密度 $q(\boldsymbol{x}_n \mid \boldsymbol{x}_{n-1}, z_n)$ 解析表达式。似然函数是基于观测噪声密度的,通常假定为高斯密度,然而当观测方程为高度非线性时,这种假设就不成立,如第 4.4 节给出的 DIFAR 浮标案例中的情况;在动态转移方程中,预测密度是基于噪声的,可以是高斯、非高斯或者完全未知的;重要性密度通常为某种已知的分析密度,可以很容易地进行采样。但要注意,$q(\boldsymbol{x}_n \mid \boldsymbol{x}_{n-1}, z_n)$ 是以当前的观测值为条件的,因此在非线性高斯卡尔曼滤波器的输出中,重要性密度通常被当作后验密度,应用于每个粒子当中,其中更新方程利用最近的观测值。序贯重要性采样粒子滤波器的这三种密度非常重要,必须认真选取。

序贯重要性采样粒子滤波器的第二个问题是粒子的发散,注意只有在初始化步骤才生成粒子。尽管第 2 步看起来是一个重采样步骤,而实际上,通常使用非线性高斯卡尔曼滤波器作为重要性密度,对每个粒子分别处理。在这种情况下,原来的粒子只是通过了滤波器而没有进行重采样。因此,最初的粒子在每次迭代中传播,在初始化步骤生成粒子后,没有新粒子生成。因此,这类蒙特卡洛

估计方法被称为粒子滤波器，因为粒子的轨迹代表通过状态空间的时间路径。生成粒子的过程是由某些初始重要性密度进行采样的过程，之后粒子就在序贯重要性采样过程中一次次传播，每一次迭代都会加入噪声。这导致基于离散后验密度的粒子传播与真实后验密度的粒子传播之间存在偏差，除少数主要的粒子之外，多数粒子的权重降为零。这就是通常所说的 SIS 退化问题。在每次粒子传播后，用所有粒子的权重来进行矩估计，但权值（重要性）足够大，这样在离散求和中起作用的粒子越来越少，增大了矩估计的方差。为降低这种增长对估计方差所造成的影响，提出了很多 SIS 粒子滤波方法。

还应指出的是，重要性密度的初始化 $q(\boldsymbol{x}_0)$ 甚至不需要与 SIS 第 1 步中用到的重要性密度 $q(\boldsymbol{x}_n|\boldsymbol{x}_{n-1},\boldsymbol{z}_n)$ 相同。通常希望初始化密度 $q(\boldsymbol{x}_0)$ 尽可能无信息，以包含后面所有可能区域的蒙特卡洛采样。但通过最新的观测值 $\boldsymbol{z}_n$，可以使 $q(\boldsymbol{x}_n|\boldsymbol{x}_{n-1},\boldsymbol{z}_n)$ 高度局域化。

16.2 序贯重要性采样粒子滤波器的重采样和正则化

SIS 样本退化问题可以通过重采样过程来缓解。重采样的基本思想是消除低权重的状态空间粒子，复制高权重粒子来替换掉低权重粒子。在这一过程中，把权重看成是一维离散概率分布，用与累积分布权重的逆相关联的逆变换方法[1,2]来对粒子进行重采样。

然而，这将造成新的问题——由于生成的样本集合中包含给定权重的多个重复粒子，会失去粒子的多样性。有文献将这种问题称为样本贫化。在严重的情况下，所有的粒子都迁移到一个采样点。为了矫正由重采样导致的样本贫化，在每次重采样过程之后，可以用粒子密度的核密度估计来对粒子进行重采样。在这一过程中，每个新粒子（蒙特卡洛样本）都是由重采样粒子中基于均匀分布选出的，并基于局部核函数对采样点进行微小移动。这个过程往往集中在概率最高区域的粒子，将它们以随机的方式分开。这种减轻样本贫化的方法被称为正则化[3]，并构成了采样与移动过程中的移动部分。

除文献[4-8]讨论的方法之外，还有一些重采样和移动的方法，包括马尔可夫链蒙特卡洛（MCMC）采样方法和以 MCMC 类似的 Gibbs 采样方法。MCMC 采用 *Metropolis-Hastings* 接受算法取代正则化步骤，这些方法对于实时递归跟踪应用来说计算量太大，这里不会予以考虑。然而，在某些情况下后验密度有大的尾概率，例如 α 稳定分布情况的 Levy 分布，选择适当的重要性密度非常困难，因此标准 SIS 粒子滤波方法可能会失败。在这种情况下，粒子滤波器中引入 MCMC 方法，提供了另一种方法用来构建高效的高维建议分布。对于 MCMC 粒子滤波的讨论可参阅文献[4]及其包含的参考文献。这些方法还可以应用到静态参数估计和平滑方法[5]当中，但本书对这两种应用不做深入讨论。

16.2.1 逆变换方法

首先查看图 2-7 中的一维累积分布函数 $P(x)$，找到解析高斯分布$\mathcal{N}(-5,6^2)$的值。如图所示，与解析 CDF 相关的累积概率在区间(0,1]内均匀分布。如果对 CDF 求逆，可得 $P^{-1}(u)$，如图 16-1 所示。如果在均匀分布(0,1]中生成样本(u_i)，并利用它选择一个沿垂直的 x 轴(x_i)方向上的值，在这个例子中，将由$\mathcal{N}(-5,6^2)$生成一个样本。注意，在这个例子当中，只在 $-5-13\leqslant x\leqslant -5+13$ 范围之外选取很少的样本。

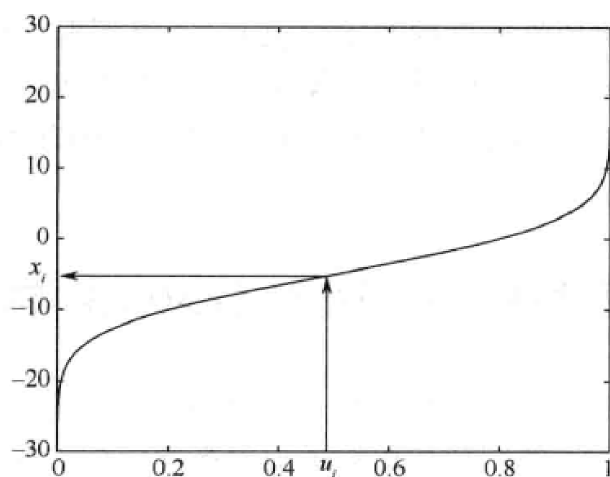


图 16-1　高斯累积分布的逆

通常，累积分布及其逆有如下属性[1,2,9]：

(1) $P^{-1}(U)$单调递增；

(2) $P(P^{-1}(U)\geqslant u$ 且 $P(P^{-1}(X)\leqslant x$；

(3) $u\leqslant P(X)$当且仅当 $P^{-1}(U)\leqslant x$；

(4) P^{-1}为左连续。

定理 16.1　如果 $u\sim U(0,1)$，那么 $x\sim P^{-1}(U)=P(X)$

证明：$P(X\leqslant x)=P(X\leqslant P^{-1}(U))=P(P^{-1}(u))=u=P(X)$

这个一般定理仅适用于解析和经验一维分布，只有少数特殊情况下，才能应用到多维分布中。用一个简单的例子来说明这个逆分布采样方法。首先利用第 2 章给出的方法，用 MATLAB 由$\mathcal{N}(-5,6^2)$生成 50000 个样本。然后，由一个经验 PDF 生成一个 13 条的直方图，并用样本的总数进行规范化。条与条之间的累加和构成了 CDF。图 16-2 由 $P(x)$与 x 及 x 与 $P^{-1}(u)$之间的关系构成。注意，对于这个简单的例子，重采样的过程只能生成 7 个 x 值。

有很多数值方法可实现生成样本的逆运算。这里只考虑其中 3 个：随机采样、分层采样和等距采样。然而，研究表明，“最优”结果可以由分层或等距采样[10]得到。

对于随机重采样，由均匀分布 $U(0,1)$生成样本$\{u_i,i=1,\cdots,n\}$，然后选取

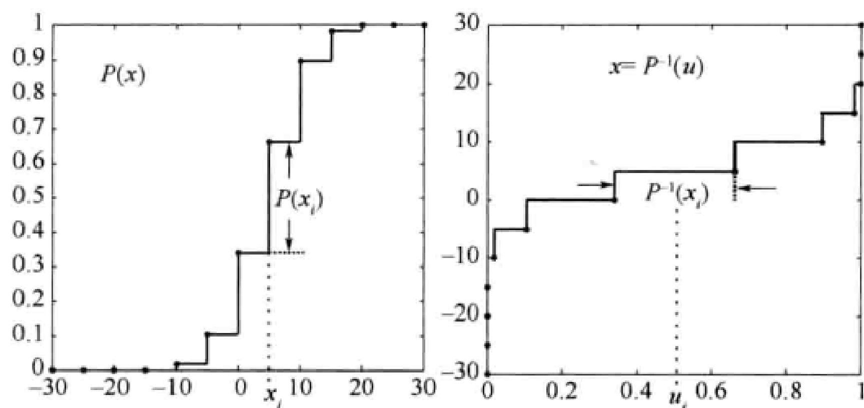


图 16－2 高斯 CDF 及其逆的对比

满足 $x_i = P^{-1}(u_i)$ 的值 $\{x_i, i=1,\cdots,n\}$。列表 16－1 给出了实现这一功能的简单的 MATLAB 源代码。

列表 16－1 随机采样 MATLAB 程序片段

```
% Random Resampling
u = cumprod(rand(1,Ns)./[Ns:-1:1]);
u = u(N:-1:1);
wc = cumsum(w);
label = nan(1,Ns);
k = 1;

for i = 1:Ns
    while(wc(k) < u(i))
        k = k + 1;
    end
    label(i) = k;
end

% Resample particles
particles = particles(:,label);
w = ones(1,Ns)./Ns;
```

在分层重采样当中，基于调查抽样的思想，将区间 $(0,1]$ 预分割为 n 个均匀的伴随集或层 $(0,1] = (0,1/n] \cup (1/n,2/n) \cup \cdots \cup (n-1/n,1]$。这样，每一层都生成了单独的均匀离散数值，例如第 i 层，有 $u_i \sim U(i/n, i+1/n)$。因此可设置 $x_i = P^{-1}(u_i)$。列表 16－2 给出了实现上述功能的 MATLAB 代码。

列表 16－2 分层采样 MATLAB 代码片段

```
% Strati fied Resampling
u = ([0:Ns-1] + rand(1,Ns)/Ns;
wc = cumsum(w);
```

```
label = nan(1,Ns);
k = 1;

for i = 1:Ns
    while(wc(k) < u(i))
        k = k + 1;
    end
    label(i) = k;
end

% Resample particles
particles = particles(:,label);
w = ones(1,Ns)./Ns;
```

最后,在等距采样中,均匀区域被再次细分为 n 个均匀的伴随层,与分层采样相同,但对于所有层只有一个单独的随机数绘制。列表 16－3 给出实现上述算法的 MATLAB 代码。

列表 16－3 等距采样 MATLAB 代码片段

```
% Systematic Resampling
u = ([0:Ns-1] + rand(1,Ns))/Ns;
wc = cumsum(w);
label = nan(1,Ns);
k = 1;

for i = 1:Ns
    while(wc(k) < u(i))
        k = k + 1;
    end
    label(i) = k;
end

% Resample particles
particles = particles(:,label);
w = ones(1,Ns)./Ns;
```

在本书中所有需要进行粒子滤波器重采样的案例分析中,均采用等距重采样。

16.2.2 带重采样的序贯重要性采样粒子滤波器

可以使用逆变换方法在 SIS 粒子滤波器中进行重采样。对于式(16－11)中

给出的离散分布 $p(\boldsymbol{x}_n|z_{1:n})$，首先形成累积分布 $p(\boldsymbol{x}_n|z_{1:n})$ 的离散逼近。由式(2－80)中给出的累积分布的定义，得

$$P(\boldsymbol{x}_n^{(j)}|\boldsymbol{z}_{1:n}) = \Pr\{\boldsymbol{y}_n|\boldsymbol{z}_{1:n} \leqslant \boldsymbol{x}_n^{(j)}|\boldsymbol{z}_{1:n}\} = \int_{-\infty}^{\boldsymbol{x}_n^{(j)}} p(\boldsymbol{y}_n|\boldsymbol{z}_{1:n})\mathrm{d}\boldsymbol{y}_n$$

$$= \int_{-\infty}^{\boldsymbol{x}_n^{(j)}} \sum_{i=1}^{N_s} w_n^{(i)}\delta(\boldsymbol{y}_n - \boldsymbol{y}_n^{(i)})\mathrm{d}\boldsymbol{y}_n = \sum_{i=1}^{j} w_n^{(i)} \tag{16-15}$$

这样，离散分布 $p(\boldsymbol{x}_n|z_{1:n})$ 的每个粒子对 CDF 贡献其权值，主要由于 CDF 是由 PDF 通过积分得的，这是一个非常有趣的结论。第 2 章提出了一种多维 CDF，这种 CDF 的分量包含的维数与 PDF 相同。对于离散 $p(\boldsymbol{x}_n|z_{1:n})$，尽管状态矢量 $\boldsymbol{x}_n$ 和 $p(\boldsymbol{x}_n|z_{1:n})$ 是多维的，也可以仅由权值构造一维 CDF。因此可以利用逆变换方法对离散多维 $p(\boldsymbol{x}_n|z_{1:n})$ 进行重采样。

这些粒子滤波重采样步骤通过修改加权近似密度，消去重要性权值低的粒子，生成未加权密度。低权值粒子的去除是通过复制一个高重要性权值粒子来取代低权值粒子，以保持粒子的数量不变。因此有

$$p(\boldsymbol{x}_n|\boldsymbol{z}_{1:n}) = \sum_{i=1}^{N_s} w_n^{(i)}\delta(\boldsymbol{x}_n - \boldsymbol{x}_n^{(i)}) \tag{16-16}$$

被替换为

$$p(\boldsymbol{x}_n^*|\boldsymbol{z}_{1:n}) = \sum_{i=1}^{N_s} \frac{1}{N_s}\delta(\boldsymbol{x}_n - \boldsymbol{x}_n^*) = \sum_{i=1}^{N_s} \frac{n_i}{N_s}\delta(\boldsymbol{x}_n - \boldsymbol{x}_n^{(i)}) \tag{16-17}$$

这里 n_i 是粒子 $\{\boldsymbol{x}_n^*\}$ 的重采样集合中粒子 $\boldsymbol{x}_n^{(i)}$ 副本的数目。参见文献[10]对这些以及其他 SIS 粒子滤波重采样方法的完整评价。在 SIS 粒子滤波器的许多应用中，由扩散迁移引起的离散粒子密度远离真实密度的粒子退化情况并不严重，因此不需要在每次迭代中都进行重采样。对于特殊应用来说，合适的退化衡量指标是有效采样大小 N_{eff}，可以由文献[11]进行估算，即

$$N_{\mathrm{eff}} = \frac{1}{\sum_{i=1}^{N_s}(w_n^{(i)})^2} \tag{16-18}$$

如果所有的权值都相等，即 $\omega_n^{(i)} = 1/N_s$，那么有 $N_{\mathrm{eff}} = N_s$。然而，如果粒子退化非常严重，例如对于粒子 i，有 $\omega_n^{(i)} = 1$；而对于所有其它粒子，权值均为零，此时 $N_{\mathrm{eff}} = 1$。因此 $N_{\mathrm{eff}} \ll N_s$ 表示退化严重，而 $N_{\mathrm{eff}} \sim N_s$ 表示退化很微小。

表 16－2 给出了重采样 SIS 粒子滤波器过程。

表 16－2　带重采样的序贯重要性采样粒子滤波器

1. 滤波器初始化	
（1）初始化状态矢量样本；	$\boldsymbol{x}_0^{(i)} \sim q(\boldsymbol{x}_0) \quad i = 1, \cdots, N_s$
（2）初始化权值	$w_0^{(i)} = \frac{1}{N_s}$

（续）

2. 序贯重要性采样 （1）选取新样本； （2）生成非规范化的重要性权值； （3）规范化权值	$x_n^{(i)} \sim q(x_n \mid x_{n-1}^{(i)}, z_n), \forall i$ $\tilde{w}_n^{(i)} = w_{n-1}^{(i)} \cdot \frac{p(z_n \mid x_n^{(i)}) p(x_n^{(i)} \mid x_{n-1}^{(i)})}{q(x_n^{(i)} \mid x_{n-1}^{(i)}, z_n)}$ $w_n^{(i)} = \tilde{w}_n^{(i)} / \sum_{i=1}^{N_s} \tilde{w}_n^{(i)}$
3. 计算 N_{eff}	$N_{\text{eff}} = \frac{1}{\sum_{i=1}^{N_s} (w_n^{(i)})^2}$
4. If $N_{\text{eff}} \ll N_s$ （1）重采样； （2）重置权值 end if	利用 MATLAB 重采样例程进行重采样 $w_n^{(i)} = 1/N_s$
5. 如需要，计算矩 （1）一阶矩：状态矢量均值； （2）二阶矩：协方差	$\hat{x}_n = \sum_{i=1}^{N_s} w_n^{(i)} x_n^{(i)}$ $P_n^{xx} = \sum_{i=1}^{N_s} w_n^{(i)} \cdot (x_n^{(i)} - \hat{x}_n)(x_n^{(i)} - \hat{x}_n)^{\mathrm{T}}$
6. 时间递增，返回步骤 2.1	$x_n^{(i)} \to x_{n-1}^{(i)}; w_n^{(i)} \to w_{n-1}^{(i)}$

16.2.3 正则化

重采样之后，在结果粒子的样本集合中，对于任意给定的权值，都包含很多重复的粒子。在文献中这种问题被称为样本贫化。为修正由重采样所导致的样本贫化，在每次重采样过程之后，可以利用粒子密度的核密度估计来对粒子进行第二次重采样。在这个过程中，每个新粒子（蒙特卡洛样本）都是在原始重采样粒子中均匀分布选取，然后基于局部核函数，对采样点进行微小的移动（抖动）。这个正则化过程往往集中在粒子概率最高的区域，用随机的方式把它们分开。图 16－3 给出这一过程的完整结果的示例。

图 16－3 中给出了后验密度和重要性密度。样本取自重要性密度，其位置沿 x 轴方向，用编号来描述。根据比率 $\omega_n^{(i)} = p(x_n^{(i)})/q(x_n^{(i)})$，为每个粒子分配一个非规范化权值，图 16－3 最上面一行编号的半径表示权值的相对大小。

这些数字占位符用来表示继承关系，即每一步结束时各个粒子的最终位置。第二行表示粒子重采样后，所有的粒子都有相等的权值，低权值粒子变为现有高权值样本的副本。第三行表示样本正则化以后，即在重采样后，移动粒子，使得没有两个粒子具有完全相同的 x 值。

表 16－3 给出了多维状态矢量的重采样并正则化（移动）SIS 粒子滤波器过程。注意，经验协方差矩阵 D_s 的计算必须在重采样步骤之前进行，因此是 $x_n^{(i)}$

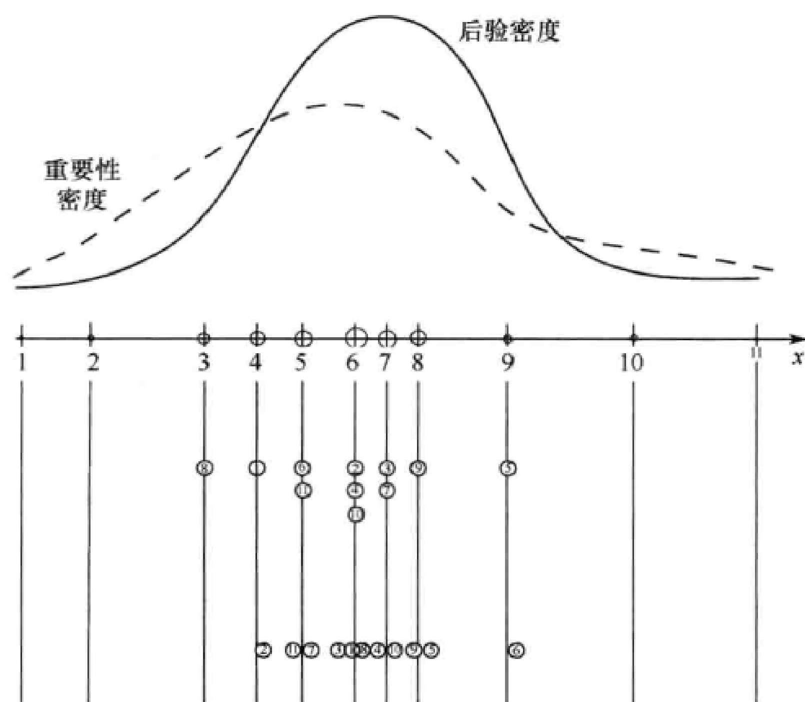


图 16－3　重采样原理

上面：从 q（虚线）中选取样本，用半径与规范化权重成比例的编号来表示相关的规范化重要性权值（p 对应的目标密度用实线来标绘）；

中间：在重采样之后，所有的粒子具有相同的重要性权值，其中一些已经复制过；

下面：样本已进行了规范化，利用核密度估计器进行重采样和移动。

和 $w_n^{(i)}$ 的函数。因为正则化步骤由后验密度进行重采样，所有粒子在矩积分估计总和中有着相等的重要性，所以将所有的权值重置为 $w_n^{(i)}=1/N$。在每一次迭代之后，可在时间递增之前进行矩的计算。

可以看出，重采样、正则化的 SIS 粒子滤波仍然无法应用于跟踪问题。在确定合适的重要性分布 $q(\boldsymbol{x}_n|\boldsymbol{x}_{n-1},\boldsymbol{z}_n)$ 以及似然函数 $p(\boldsymbol{z}_n|\boldsymbol{x}_n)$ 的解析形式上，还存在很多问题。

表 16－3　序贯重要性采样粒子滤波器正则化

1. 滤波器初始化	
（1）初始化状态矢量样本；	$\boldsymbol{x}_0^{(i)} \sim q(\boldsymbol{x}_0) \quad i=1,\cdots,N_s$
（2）初始化权值	$w_0^{(i)}=\dfrac{1}{N_s}$
2. 序贯重要性采样	
（1）选取新样本；	$\boldsymbol{x}_n^{(i)} \sim q(\boldsymbol{x}_n \mid \boldsymbol{x}_{n-1}^{(i)},\boldsymbol{z}_n),\ \forall i$
（2）生成非规范化的重要性权值；	$\tilde{w}_n^{(i)}=w_{n-1}^{(i)}\cdot\dfrac{p(\boldsymbol{z}_n \mid \boldsymbol{x}_n^{(i)})p(\boldsymbol{x}_n^{(i)} \mid \boldsymbol{x}_{n-1}^{(i)})}{q(\boldsymbol{x}_n^{(i)} \mid \boldsymbol{x}_{n-1}^{(i)},\boldsymbol{z}_n)}$
（3）规范化权值	$w_n^{(i)}=\tilde{w}_n^{(i)}/\sum_{i=1}^{N_s}\tilde{w}_n^{(i)}$

（续）

3. 计算 N_{eff}	$N_{\mathrm{eff}} = \dfrac{1}{\sum_{i=1}^{N_s} (w_n^{(i)})^2}$
4. If $N_{\mathrm{eff}} \ll N_s$	
（1）计算经验平均值	$\boldsymbol{\mu}_s = \dfrac{1}{N_s} \sum_{i=1}^{N_s} \boldsymbol{x}_n^{(i)}$
（2）计算经验协方差	$\boldsymbol{\Sigma}_s = \dfrac{1}{N_s} \sum_{i=1}^{N_s} (\boldsymbol{x}_n^{(i)} - \boldsymbol{\mu}_s) \cdot (\boldsymbol{x}_n^{(i)} - \boldsymbol{\mu}_s)^{\mathrm{T}}$ $\boldsymbol{\Sigma}_s = \boldsymbol{D}_s \boldsymbol{D}_s^{\mathrm{T}}$
（3）粒子重采样	利用 MATLAB 重采样例程进行重采样
（4）粒子正则化	利用表 15 - 3 中的步骤，实现粒子正则化
（5）时间递增，返回步骤 2.1	$\boldsymbol{x}_n^{(i)} \to \boldsymbol{x}_{n-1}^{(i)}; w_n^{(i)} = 1/N$
Else	
时间递增，返回步骤 2.1	$\boldsymbol{x}_n^{(i)} \to \boldsymbol{x}_{n-1}^{(i)}; w_n^{(i)} \to w_{n-1}^{(i)}$
end if	

16.3 自举粒子滤波器

自举粒子滤波器（BPF）在文献[12]中首次提出，因为最容易实现，也是使用最为广泛的重采样 SIS 粒子滤波器。在自举粒子滤波器中，选择转移密度作为重要性密度，即

$$q(\boldsymbol{x}_n \mid \boldsymbol{x}_{n-1}, \boldsymbol{z}_n) = p(\boldsymbol{x}_n \mid \boldsymbol{x}_{n-1}) \tag{16-19}$$

由于选取了重要性密度，权重更新方程（16 - 12）变成

$$\widetilde{w}_n^{(i)} = w_{n-1}^{(i)} p(\boldsymbol{z}_n \mid \boldsymbol{x}_n^{(i)}) \tag{16-20}$$

自举粒子滤波器的特征是：权重的增量不依赖于过去的粒子轨迹，只取决于观测值 $p(\boldsymbol{z}_n | \boldsymbol{x}_n)$ 的条件似然函数。自举粒子滤波器的采样非常简单，用状态转移方程（5 - 1）转移上一时间步中的粒子，生成新的粒子。然后是重采样和移动步骤，表 16 - 4 给出了完整的 BPF 过程。同样，分布的一阶和二阶矩可以在每个时间步结束时进行计算。

需要强调一点，表 16 - 4 中，在状态协方差计算时缺少噪声协方差矩阵 $\boldsymbol{Q}$。这里并不需要计算 $\boldsymbol{Q}$，因为噪声项 $\boldsymbol{v}_n^{(i)} \sim p(\boldsymbol{v}_n)$ 的抽样已经包含在状态矢量的转移方程中。因此，这里也需要动态噪声密度的信息。因为最初的假设是针对后验密度未知的情况，那么假定动态噪声的解析形式违背了最初的假设。绕过这一问题的常用方法是忽略噪声项，在粒子转移过程中不添加噪声。

还有就是需要找到似然函数 $p(\boldsymbol{z}_n | \boldsymbol{x}_n)$ 的解析表达式。如前所述，$p(\boldsymbol{z}_n | \boldsymbol{x}_n)$ 通常被认为是高斯函数。但对于高度非线性观测方程，观测噪声可以是非高斯

的，例如在 DIFAR 实例当中。

用自举粒子滤波器解决实际问题时，主要问题是当观测噪声很小的时候，$p(\boldsymbol{z}_n|\boldsymbol{x}_n)$是一个非常窄的函数。这时，转移粒子的$p(\boldsymbol{z}_n|\boldsymbol{x}_n^{(i)})$值非常小，导致相应的更新权值也非常小。在这种情况下，需要在每一次更新时进行重采样和移动步骤。另一方面，如果假设$p(\boldsymbol{z}_n|\boldsymbol{x}_n)$为高斯分布（实际上并不是），在观测值当中会有无信息的异常值，例如所有粒子$p(\boldsymbol{z}_n|\boldsymbol{x}_n^{(i)})$值都为零将导致滤波器不稳定。

表 16－4　带重采样和正则化的自举 SIS 粒子滤波器

1. 滤波器初始化	
（1）初始化状态矢量样本；	$\boldsymbol{x}_0^{(i)} \sim q(\boldsymbol{x}_0) \quad i=1,\cdots,N_s$
（2）初始化权值	$w_0^{(i)} = \frac{1}{N_s}$
2. 序贯重要性采样	
（1）选取新样本；	$\boldsymbol{x}_n^{(i)} \sim p(\boldsymbol{v}_n)$ $\boldsymbol{x}_n^{(i)} = \boldsymbol{f}(\boldsymbol{x}_{n-1}^{(i)}) + \boldsymbol{v}_n^{(i)}, \forall i$
（2）生成非规范化的重要性权值；	$\widehat{w}_n^{(i)} = w_{n-1}^{(i)} p(\boldsymbol{z}_n \mid \boldsymbol{x}_n^{(i)})$
（3）规范化权值	$w_n^{(i)} = \widehat{w}_n^{(i)} / \sum_{i=1}^{N_s} \widehat{w}_n^{(i)}$
3. 计算 N_{eff}	$N_{\text{eff}} = \frac{1}{\sum_{i=1}^{N_s} (w_n^{(i)})^2}$
4. If $N_{\text{eff}} \ll N_s$	
（1）计算经验平均值；	$\boldsymbol{\mu}_s = \frac{1}{N_s} \sum_{i=1}^{N_s} \boldsymbol{x}_n^{(i)}$
（2）计算经验协方差；	$\boldsymbol{\Sigma}_s = \frac{1}{N_s} \sum_{i=1}^{N_s} (\boldsymbol{x}_n^{(i)} - \boldsymbol{\mu}_s) \cdot (\boldsymbol{x}_n^{(i)} - \boldsymbol{\mu}_s)^{\mathrm{T}}$ $\boldsymbol{\Sigma}_s = \boldsymbol{D}_s \boldsymbol{D}_s^{\mathrm{T}}$
（3）粒子重采样；	利用 MATLAB 重采样例程进行重采样
（4）重采样粒子正则化；	利用表 15－3 中的步骤，实现粒子正则化
（5）时间递增，返回步骤 2.1	$\boldsymbol{x}_n^{(i)} \to \boldsymbol{x}_{n-1}^{(i)}; w_n^{(i)} = 1/N$
Else	
时间递增，返回步骤 2.1	$\boldsymbol{x}_n^{(i)} \to \boldsymbol{x}_{n-1}^{(i)}; w_n^{(i)} \to w_{n-1}^{(i)}$
end if	

16.3.1　BPF 在 DIFAR 浮标跟踪中的应用

第 4 章式（4－34）给出了 DIFAR 浮标的似然函数$p(\theta_{n,m}|\vartheta_m)$，用来表示$t_n$时刻第$m$个 DIFAR 元素的方位测量值$\theta_{n,m}$，以整个空间的可能目标方位$\{-\pi \leqslant \vartheta_m \leqslant \pi\}$为条件，作为浮标输入的信号信噪比。高信噪比信号的规范化似然函数

为高斯的，而低信噪比的接近于均匀分布，这并不出人意料。因为在信号缺失（低信噪比）的情况下，浮标的输入是在$\{-\pi\leqslant\theta_{n,m}\leqslant\pi\}$区间呈均匀分布的噪声。假设噪声独立于浮标，可以将浮标场似然函数写成单个浮标函数的乘积，并将其应用到表16－4 自举粒子滤波器过程中的第2步。

为初始化自举粒子滤波器，使用与7.4.3中相同的步骤。在计算初始目标状态$\boldsymbol{x}_0$估计之后，生成一组初始化粒子$\{\boldsymbol{x}_0^{(i)},i=1,\cdots,N_s\}$，有

$$\boldsymbol{x}_0^{(i)}=\boldsymbol{x}_0+\boldsymbol{v}_n^{(i)} \tag{16-21}$$

其中$\boldsymbol{v}_n^{(i)}\sim\mathcal{N}(\boldsymbol{0},\boldsymbol{Q})$。顺序执行表16－4中其余的步骤，得如图16－4所示的估计目标轨迹与真实轨迹的对比。这里设信噪比为20dB，初始距离为4n mile，速度为30kn（实际的初始距离和速度大约分别为3.5n mile和30kn）。对于这个例子，设粒子总数为1000。

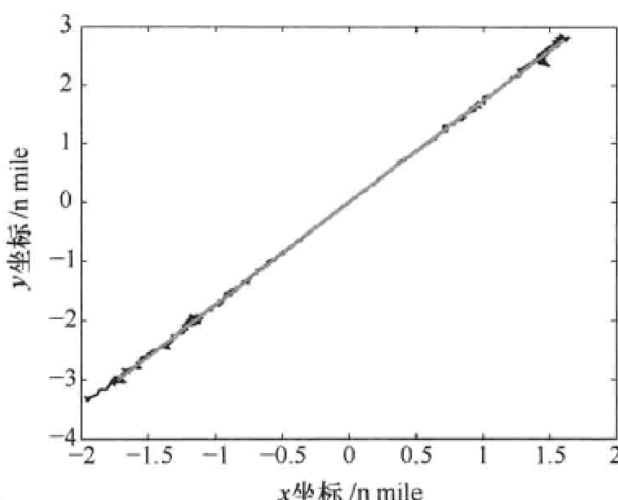


图16－4　利用自举粒子滤波器实现DIFAR浮标场的目标跟踪

图16－5表明6种不同输入信号信噪比（从20dB降至0dB，步长为5dB）的情况下，自举粒子滤波器的跟踪输出结果与从前的DIFAR跟踪估计中的初始化相同。在这个例子中，BPF使用了3000个粒子。与图7－3给出的类似的EKF标绘图相比，可发现在所有信噪比下，EKF的跟踪结果都比BPF更精确。而且，在每种信噪比下，100次BPF蒙特卡洛运行都出现了发散轨迹。在后面将讨论产生BPF性能缺陷的一些原因。

与高斯卡尔曼滤波器不同，BPF的性能对距离的初值非常敏感。如果初始距离大于5n mile，BPF航迹将会发散。增加粒子的数目会在一定程度上缓解这一问题，但可以发现，由于其对初值的敏感，BPF实际很难使用，BPF似乎对于初始速度的变化不太敏感。本书试图通过选择无信息的均匀分布距离和速度来初始化粒子，距离范围2n mile～5.9n mile，增量0.1n mile，速度范围24kn～30kn，增量为0.5kn，但多数跟踪估计都是高度发散的。虽然在文献中很少提及，这种初值敏感性使BPF难以应用于真实场景中，除非有另一种跟踪方法来生成“好

的"初值。此外,必须注意的是,对 BPF 算法,其重要性密度不考虑当前观测值,这使 BPF 成为次优滤波器选择。虽然 BPF 算法很容易实现,但相比一些下面将提出的粒子滤波方法来说,它有严重的缺陷。

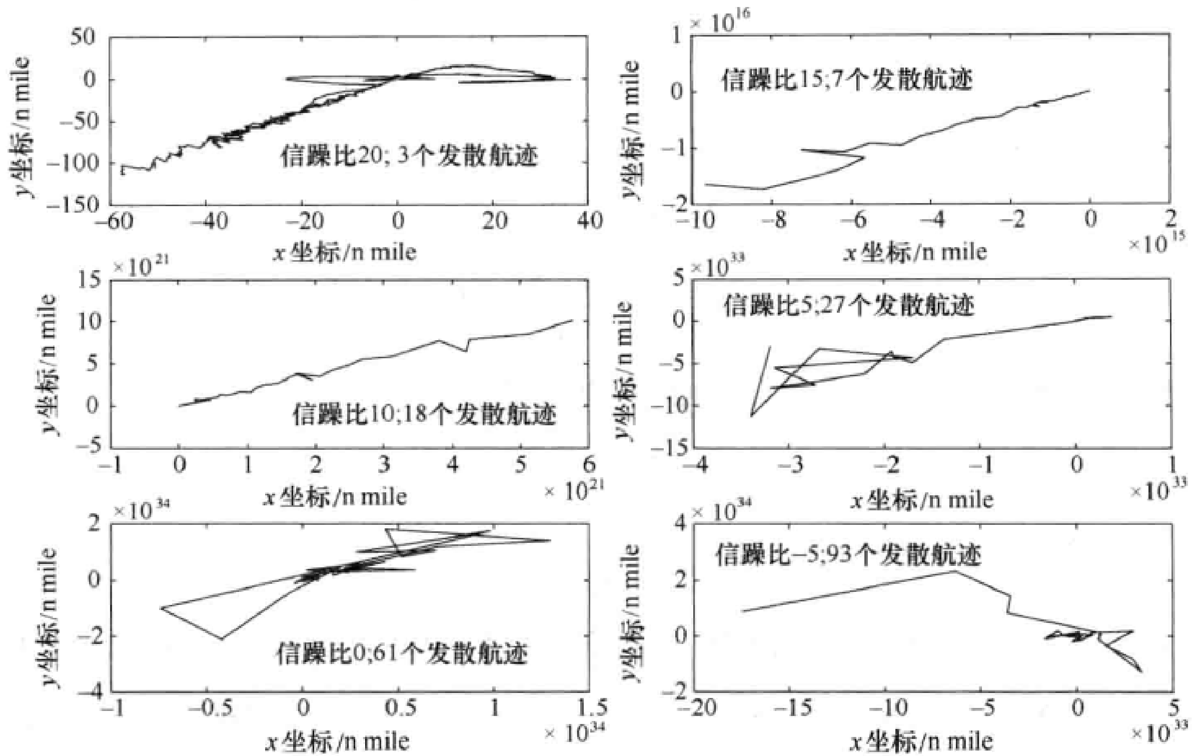


图 16－5　六种信噪比下 BPF 跟踪器输出结果比较

16.4　最优序贯重要性采样粒子滤波器

为克服 BPF 的上述问题,需要合理地选择重要性密度。因此,在设计粒子滤波器时,重要性密度的选择是最关键的问题。重要性密度的最优选择是使权值的方差最小。权值的均值和方差可表示为

$$\boldsymbol{\varepsilon}_{q(\boldsymbol{x}_n^{(i)}|\boldsymbol{x}_{n-1}^{(i)},z_n)}\{w_n^{*(i)}\} = \int w_n^{*(i)} q(\boldsymbol{x}_n^{(i)} \mid \boldsymbol{x}_{n-1}^{(i)}, z_n)\mathrm{d}\boldsymbol{x}_n \tag{16-22}$$

$$\mathrm{Var}_{q(\boldsymbol{x}_n^{(i)}|\boldsymbol{x}_{n-1}^{(i)},z_n)}\{w_n^{*(i)}\} = \boldsymbol{\varepsilon}_{q(\boldsymbol{x}_n^{(i)}|\boldsymbol{x}_{n-1}^{(i)},z_n)}\{[w_n^{*(i)}]^2\} - [\boldsymbol{\varepsilon}_{q(\boldsymbol{x}_n^{(i)}|\boldsymbol{x}_{n-1}^{(i)}),z_n})\{w_n^{*(i)}\}]^2 \tag{16-23}$$

如果首先检查均值方程(16－22),将其代入式(16－12),得 w_n^{*i} 的递归方程为

$$\begin{aligned}\boldsymbol{\varepsilon}_{q(\boldsymbol{x}_n^{(i)}|\boldsymbol{x}_{n-1}^{(i)},z_n)}\{w^{*(i)}{}_n\} &= w_{n-1}^{*(i)}\int p(z_n \mid \boldsymbol{x}_n)p(\boldsymbol{x}_n \mid \boldsymbol{x}_{n-1}^{(i)})\mathrm{d}\boldsymbol{x}_n \\ &= \widetilde{w}_{n-1}^{*(i)} p(z_n \mid \boldsymbol{x}_{n-1}^{(i)})\end{aligned} \tag{16-24}$$

对方差方程(16－23)的第一项采用相同的步骤,并利用式(16－24)计算

式(16 - 23)的第二项,得

$$\mathrm{Var}_{q(\boldsymbol{x}_n^{(i)}\mid\boldsymbol{x}_{n-1}^{(i)},z_n)}\{w_n^{*(i)}\} = [w_{n-1}^{*(i)}]^2\left\{\int\frac{[p(z_n\mid\boldsymbol{x}_n)p(\boldsymbol{x}_n\mid\boldsymbol{x}_{n-1}^{(i)})]}{q(\boldsymbol{x}_n\mid\boldsymbol{x}_{n-1}^{(i)},z_n)}\mathrm{d}\boldsymbol{x}_n - [p(z_n\mid\boldsymbol{x}_{n-1}^{(i)})]^2\right\} \tag{16-25}$$

根据贝叶斯定律,如果满足下式,式(16 - 25)可化简为零[13]

$$q(\boldsymbol{x}_n\mid\boldsymbol{x}_{n-1}^{(i)},z_n) = p(\boldsymbol{x}_n\mid\boldsymbol{x}_{n-1}^{(i)},z_n) \tag{16-26}$$

$$= \frac{p(z_n\mid\boldsymbol{x}_n,\boldsymbol{x}_{n-1}^{(i)})p(\boldsymbol{x}_n\mid\boldsymbol{x}_{n-1}^{(i)})}{p(z_n\mid\boldsymbol{x}_{n-1}^{(i)})} \tag{16-27}$$

这样,重要性权值更新方程(16 - 12)的表达式变为

$$\widetilde{w}_n^{(i)} = w_{n-1}^{(i)}\frac{p(z_n\mid\boldsymbol{x}_n^{(i)})p(\boldsymbol{x}_n^{(i)}\mid\boldsymbol{x}_{n-1}^{(i)})}{p(\boldsymbol{x}_n^{(i)}\mid\boldsymbol{x}_{n-1}^{(i)},z_n)} \tag{16-28}$$

$$= w_{n-1}^{(i)}p(z_n\mid\boldsymbol{x}_{n-1}^{(i)}) \tag{16-29}$$

将 BPF 权值方程(16 - 20)与 OPF 权值方程(16 - 29)进行比较,可以看到,OPF 重要性权值 $\widetilde{\omega}_n^{(i)}$ 不再依赖于 $\boldsymbol{x}_n^{(i)}$。因此,t_n 时刻的重要性权值可以在粒子传播至 t_n 时刻之前计算得出。

使用最优重要性密度有两个缺点[14]。首先,必须可以由重要性密度 $p(\boldsymbol{x}\mid\boldsymbol{x}_{n-1}^{(i)},z_n)$采样;其次,必须能将 $p(z_n\mid\boldsymbol{x}_{n-1}^{(i)})$计算到一个规范化常数,其中 $p(z_n\mid\boldsymbol{x}_{n-1}^{(i)})$由积分求得,即

$$p(z_n\mid\boldsymbol{x}_{n-1}^{(i)}) = \int p(z_n\mid\boldsymbol{x}_n)p(\boldsymbol{x}_n\mid\boldsymbol{x}_{n-1}^{(i)})\mathrm{d}\boldsymbol{x}_n \tag{16-30}$$

通常情况下,这两个任务都是非常困难的,然而,也有一些有趣的特例可以应用最优 SIS 粒子滤波器。如文献[3,15]中指出:当 $\boldsymbol{x}_n$ 是有限集合的成员时,第一种特殊情况出现,式(16 - 30)中的积分将变成求和,$p(\boldsymbol{x}_n\mid\boldsymbol{x}_{n-1}^{i},z_n)$采样变成可能性。这种特殊情况的例子是用线性跳跃马尔可夫过程处理机动目标跟踪。下面将处理的是第二种情况,$p(\boldsymbol{x}_n\mid\boldsymbol{x}_{n-1}^{i},z_n)$是一类高斯分布的模型[13,14]。

16.4.1 高斯最优序贯重要性采样粒子滤波器

假设系统动态转移模型是非线性的,具有加性高斯噪声,而观测方程是线性的,带有加性高斯噪声。这样一个系统可以描述为

$$\boldsymbol{x}_n = \boldsymbol{f}(\boldsymbol{x}_{n-1}) + \boldsymbol{v}_{n-1} \tag{16-31}$$

$$z_n = \boldsymbol{H}_{nxn} + \boldsymbol{w}_n \tag{16-32}$$

$$\boldsymbol{v}_n \sim \mathcal{N}(\boldsymbol{0},\boldsymbol{Q}_n) \tag{16-33}$$

$$\boldsymbol{w}_n \sim \mathcal{N}(\boldsymbol{0},\boldsymbol{R}_n) \tag{16-34}$$

密度 $p(\boldsymbol{x}_n|\boldsymbol{x}_{n-1}^i)$ 和 $p(z_n|\boldsymbol{x}_n)$ 可定义为

$$p(\boldsymbol{x}_n \mid \boldsymbol{x}_{n-1}) = \mathcal{N}(\boldsymbol{x}_n;\boldsymbol{f}(\boldsymbol{x}_{n-1}),\boldsymbol{Q}_{n-1}) \tag{16-35}$$

$$p(z_n \mid \boldsymbol{x}_n) = \mathcal{N}(z_n;\boldsymbol{H}_n\boldsymbol{x}_n,\boldsymbol{R}_n) \tag{16-36}$$

由贝叶斯法则，得

$$p(\boldsymbol{x}_n \mid \boldsymbol{x}_{n-1},z_n) = \frac{p(z_n \mid \boldsymbol{x}_n)p(\boldsymbol{x}_n \mid \boldsymbol{x}_{n-1})}{p(z_n \mid \boldsymbol{x}_{n-1})} \tag{16-37}$$

由上得

$$p(z_n \mid \boldsymbol{x}_n)p(\boldsymbol{x}_n \mid \boldsymbol{x}_{n-1}) = p(\boldsymbol{x}_n \mid \boldsymbol{x}_{n-1},z_n)p(z_n \mid \boldsymbol{x}_{n-1}) \tag{16-38}$$

由于等式左边的乘积是高斯分布的乘积，那么右边的乘积一定是高斯分布的乘积。因此，令

$$p(\boldsymbol{x}_n \mid \boldsymbol{x}_{n-1},z_n) = \mathcal{N}(\boldsymbol{x}_n;\boldsymbol{a}_n,\boldsymbol{\Sigma}_n) \tag{16-39}$$

$$p(z_n \mid \boldsymbol{x}_{n-1}) = \mathcal{N}(z_n;\boldsymbol{b}_n,\boldsymbol{S}_n) \tag{16-40}$$

这样式(16－38)可以写成高斯密度项，即

$$\mathcal{N}(z_n;\boldsymbol{H}_n\boldsymbol{x}_n,\boldsymbol{R}_n)\ \mathcal{N}(\boldsymbol{x}_n;\boldsymbol{f}(\boldsymbol{x}_{n-1}),\boldsymbol{Q}_{n-1}) = \mathcal{N}(z_n;\boldsymbol{b}_n,\boldsymbol{S}_n)\ \mathcal{N}(\boldsymbol{x}_n;\boldsymbol{a}_n,\boldsymbol{\Sigma}_n) \tag{16-41}$$

解析高斯密度的指数项相等，经过代数运算，得

$$\boldsymbol{b}_n = \boldsymbol{H}_n\boldsymbol{f}(\boldsymbol{x}_{n-1}) \tag{16-42}$$

$$\boldsymbol{S}_n = \boldsymbol{H}_n\boldsymbol{Q}_{n-1}\boldsymbol{H}_n^{\mathrm{T}} + \boldsymbol{R}_n \tag{16-43}$$

$$\boldsymbol{a}_n = \boldsymbol{f}(\boldsymbol{x}_{n-1}) + \boldsymbol{\Sigma}_n\boldsymbol{H}_n^{\mathrm{T}}\boldsymbol{R}_n^{-1}(z_n - \boldsymbol{b}_n) \tag{16-44}$$

$$\boldsymbol{\Sigma}_n = \boldsymbol{Q}_{n-1} - \boldsymbol{Q}_{n-1}\boldsymbol{H}_n^{\mathrm{T}}\boldsymbol{S}_n^{-1}\boldsymbol{H}_n\boldsymbol{Q}_{n-1} \tag{16-45}$$

注意，如果 $\boldsymbol{Q}_{n-1}$、$\boldsymbol{R}_n$ 和 $\boldsymbol{H}_n$ 与时间无关，那么 $\boldsymbol{S}_n$ 和 $\boldsymbol{\Sigma}_n$ 也是时间无关的，可以离线计算。

因为这是一个粒子滤波器，可以将粒子 $\{\boldsymbol{x}_{n-1}^{(i)}, i=1,\cdots,N\}$ 插入式(16－42)和式(16－44)，所以必须令 $\boldsymbol{b}_n \to \boldsymbol{b}_n^{(i)}$ 且 $\boldsymbol{a}_n \to \boldsymbol{a}_n^{(i)}$。表 16－5 给出了 OPF 的完整步骤，符号 $\boldsymbol{x}_{n|n-1}^{(i)}$ 表示状态预测和中间的计算步骤中的预测粒子。在中间的计算步骤，用 z_n 表示 t_n 时刻的观测值。因为 SIS 粒子滤波器需要进行重采样，假设在每个时间步都需要进行重采样，也可以改成当 $N_{\mathrm{eff}} \ll N_s$ 且需要时进行重采样。注意，t_n 时刻的重要性权值是在重要性采样步骤中粒子向前传播之前、在权值更新步骤进行计算。而且，旧的粒子也可以在粒子在时间上传播之前进行重

采样，权值设为 $1/N_s$。

表 16-5 带重采样和正则化的最优 SIS 粒子滤波器

1. 滤波器初始化 (1) 初始化状态矢量样本； (2) 初始化权值	$x_0^{(i)} \sim q(x_0) \quad i=1,\cdots,N_s$ $w_0^{(i)} = \frac{1}{N_s}$
2. 状态预测	$\boldsymbol{x}_{n\|n-1}^{(i)} = \boldsymbol{f}(\boldsymbol{x}_{n-1}^{(i)})$ $\boldsymbol{b}_n^{(i)} = \boldsymbol{H}_n\boldsymbol{x}_{n\|n-1}^{(i)}$ $\boldsymbol{S}_n = \boldsymbol{H}_n\boldsymbol{Q}_{n-1}\boldsymbol{H}_n^{\mathrm{T}} + \boldsymbol{R}_n$
3. 权值更新 (1) 似然函数 (2) 权值更新 (3) 规范化权值	$p(z_n \mid \boldsymbol{x}_{n-1}^{(i)}) = \mathcal{N}(z_n;\boldsymbol{b}_n^{(i)},\boldsymbol{S}_n)$ $\widetilde{w}_n^{(i)} = w_{n-1}^{(i)}p(z_n \mid \boldsymbol{x}_{n-1}^{(i)})$ $w_n^{(i)} = \widetilde{w}_n^{(i)} / \sum_{i=1}^{N_s}\widetilde{w}_n^{(i)}$
4. 粒子重采样	利用 MATLAB 重采样例程进行重采样
5. 移动重采样粒子 重置权值	采用表 15-3 中的步骤对粒子进行正则化 $w_n^{(i)} = 1/N$
6. 重要性采样 (1) 中间计算； (2) 重要性采样	$\boldsymbol{a}_n^{(i)} = \boldsymbol{x}_{n\|n-1}^{(i)} + \boldsymbol{\Sigma}_n\boldsymbol{H}_n^{\mathrm{T}}\boldsymbol{R}_n^{-1}(z_n - \boldsymbol{b}_n^{(i)})$ $\boldsymbol{\Sigma}_n = \boldsymbol{Q}_{n-1} - \boldsymbol{Q}_{n-1}\boldsymbol{H}_n^{\mathrm{T}}\boldsymbol{S}_n^{-1}\boldsymbol{H}_n\boldsymbol{Q}_{n-1}$ $\boldsymbol{x}_n^{(i)} \sim p(\boldsymbol{x}_n \mid \boldsymbol{x}_{n-1},z_n) = \mathcal{N}(\boldsymbol{x}_n;\boldsymbol{a}_n^{(i)},\boldsymbol{\Sigma}_n)$
7. 矩计算 (1) 均值； (2) 方差	$\hat{\boldsymbol{x}}_n = \sum_i w_n^{(i)}\boldsymbol{x}_n^{(i)}$ $\boldsymbol{P}_n^{xx} = \sum_i w_n^{(i)}\boldsymbol{x}_n^{(i)}\boldsymbol{x}_n^{(i)} - \hat{\boldsymbol{x}}_n\hat{\boldsymbol{x}}_n^{\mathrm{T}}$ $\boldsymbol{x}_n^{(i)} \to \boldsymbol{x}_{n-1}^{(i)}$
8. 时间递增，返回第 2 步	

16.4.2 局部线性化的高斯最优序贯重要性采样粒子滤波器

当动态方程和观测方程均为非线性时，如果用某种方式[15]将非线性观测方程线性化，仍然可以用 OPF。考虑观测方程

$$z_n = \boldsymbol{h}(\boldsymbol{x}_n) + \boldsymbol{w}_n \tag{16-46}$$

将 $\boldsymbol{h}(\boldsymbol{x}_n)$ 展开为关于 $\boldsymbol{f}(\boldsymbol{x}_{n-1})$ 的多维泰勒级数，利用式(2-62)得

$$\boldsymbol{h}(\boldsymbol{x}_n) \approx \boldsymbol{h}(\boldsymbol{f}(\boldsymbol{x}_{n-1})) + \hat{\boldsymbol{H}}_n(\boldsymbol{x}_n - \boldsymbol{f}(\boldsymbol{x}_{n-1})) \tag{16-47}$$

雅可比矩阵 $\hat{\boldsymbol{H}}_n$ 定义为

$$\hat{\boldsymbol{H}}_n \triangleq [\boldsymbol{\nabla}_{\boldsymbol{x}_n}^{\mathrm{T}} \boldsymbol{h}(\boldsymbol{x}_n)]_{\boldsymbol{x}_n=\boldsymbol{f}(\boldsymbol{x}_{n-1})}^{\mathrm{T}} \tag{16-48}$$

这样式(16－46)变成

$$\tilde{\boldsymbol{z}}_n \approx \hat{\boldsymbol{H}}_{n}\boldsymbol{x}_n + \boldsymbol{w}_n \tag{16-49}$$

其中

$$\tilde{\boldsymbol{z}}_n \triangleq \boldsymbol{z}_n + \hat{\boldsymbol{H}}_n \boldsymbol{f}(\boldsymbol{x}_{n-1}) - \boldsymbol{h}(\boldsymbol{f}(\boldsymbol{x}_{n-1})) \tag{16-50}$$

这样，把观测方程变成在 $\boldsymbol{x}_n$ 为线性，代价是在观测方程中增加额外的项，即，$\boldsymbol{z}_n \to \tilde{\boldsymbol{z}}_n$。可以利用表 16－5 中的步骤，直接令 $\boldsymbol{H}_n \to \hat{\boldsymbol{H}}_n$ 且 $z_n \to \tilde{z}_n$，额外增加雅可比矩阵 $\hat{\boldsymbol{H}}_n$ 的计算和额外的 $\tilde{z}_n$ 项，在式(16－49)和式(16－50)中进行 $\boldsymbol{x}_n \to \boldsymbol{x}_n^{(i)}$ 和 $\boldsymbol{x}_{n-1} \to \boldsymbol{x}_{n-1}^{(i)}$ 的替换。

只要所有密度都是高斯的，可以将 OPF 应用于非线性动态方程和观测方程的任意跟踪问题。对于多数有上述特征的跟踪问题，使用 OPF 并未获得比本书第Ⅱ部分给出的所有方法更好的结果。然而，在一些实例中可能需要对跟踪解施加约束，例如将跟踪估计限制到一个预先确定的道路网格中，或限制目标状态的可操作性或速度。OPF 最适合解决这种类型的跟踪问题。因为 DIFAR 跟踪实例中对目标状态没有任何限制，本书不会将局部线性化高斯最优 SIS 粒子滤波算法应用到这个案例中。

16.5 序贯重要性采样辅助粒子滤波器

选择一个合理的 SIS 建议分布 $q(\boldsymbol{x}_n|\boldsymbol{x}_{n-1},\boldsymbol{z}_n)$，会保证将当前的观测值 z_n 并入建议步骤，这样粒子不会盲目地移入没有给出观测值的状态空间区域。显然，BPF 无法实现这一目标，因为在 $q(\boldsymbol{x}_n|\boldsymbol{x}_{n-1},\boldsymbol{z}_n) \overset{\mathrm{BPF}}{=} p(\boldsymbol{x}_n|\boldsymbol{x}_{n-1})$ 中没有考虑当前观测值。如前所述，OPF 选择 $q(\boldsymbol{x}_n|\boldsymbol{x}_{n-1},\boldsymbol{z}_n) \overset{\mathrm{OPF}}{=} p(\boldsymbol{x}_n|\boldsymbol{x}_{n-1},\boldsymbol{z}_n)$，但除了在特殊情况下，OPF 很难实现。SIS 辅助粒子滤波器(APF)是一种以次最优的方式将当前观测值加入到建议分布的粒子滤波器。

在包括 BPF 在内的标准重采样 SIS 粒子滤波中，在从重要性密度中抽取粒子和通过当前观测值更新粒子权值后，对粒子进行重采样。另一方面，在 OPF 中，重采样发生在由重要性密度中抽取粒子之前，通过对比表 16－5 与表 16－2、表 16－4 可以证实这一点。在文献[16]和第 13 章的文献[17]中，Pitt 和 Shephard 提出将 APF 作为标准 SIS 的变体，该算法的重采样步骤利用 t_n 时刻的观测值，在 t_{n-1}时刻执行。通过这种方式，APF 试图模拟最优粒子滤波器所使

用的方法。在文献[7,8]中可以找到 APF 步骤的清晰描述,本节中的部分内容也是直接引自文献[7]。

APF 的主要思想是增加权值大的("好的")粒子 $\boldsymbol{x}_{n-1}^{(i)}$,在某种意义上,这些粒子的预测依然 $p(z_n|\boldsymbol{x}_{n-1}^{(i)})$ 很大。根据贝叶斯法则(3-23),后验密度可写为

$$p(\boldsymbol{x}_n \mid z_{1:n}) \propto p(z_n \mid \boldsymbol{x}_n)p(\boldsymbol{x}_n \mid z_{1:n-1}) \tag{16-51}$$

利用 Chapman - Kolmogorov 方程(3-24),将上式变为

$$p(\boldsymbol{x}_n \mid z_{1:n}) \propto p(z_n \mid \boldsymbol{x}_n)\int p(\boldsymbol{x}_n \mid \boldsymbol{x}_{n-1})p(\boldsymbol{x}_{n-1} \mid z_{1:n-1})\mathrm{d}\boldsymbol{x}_{n-1} \tag{16-52}$$

如果有

$$p(\boldsymbol{x}_{n-1} \mid z_{1:n-1}) = \sum_{i=1}^{N_s} w_{n-1}^{(i)}\delta(\boldsymbol{x}_{n-1} - \boldsymbol{x}_{n-1}^{(i)}) \tag{16-53}$$

那么则有

$$p(\boldsymbol{x}_n \mid z_{1:n}) \propto \left[\sum_{i=1}^{N_s} w_{n-1}^{(i)}p(\boldsymbol{x}_n \mid \boldsymbol{x}_{n-1}^{(i)})\right]p(z_n \mid \boldsymbol{x}_n) \tag{16-54}$$

可将乘积 $w_{n-1}^{(i)}p(\boldsymbol{x}_n|\boldsymbol{x}_{n-1}^{(i)})$ 当作联合概率,用于计算滤波后的密度总和。

一种从经验预测密度中采样的方法是将 $\sum_i w_{n-1}^{(i)}p(\boldsymbol{x}_n|\boldsymbol{x}_{n-1}^{(i)})$ 当成一个"先验"混合密度 $\hat{p}(\boldsymbol{x}_n|z_{1:n-1})$,将其与似然度 $p(z_n|\boldsymbol{x}_n)$ 相结合,产生后验密度。样本 $\boldsymbol{x}_n^{(i)}$ 可以由 $\hat{p}(\boldsymbol{x}_n|z_{1:n-1})$ 中抽取,通过选取 $\boldsymbol{x}_n^{(i)}$ 和概率 $w_{n-1}^{(i)}$ 由 $p(\boldsymbol{x}_n|\boldsymbol{x}_{n-1}^{(i)})$ 中进行选取。

为理解 APF 方法,在采样 $q(\boldsymbol{x}_n|\boldsymbol{x}_{n-1}^{(i)},z_n)$ 之前,首先检查 SIS 更新方程(16-12),将其重写为

$$\widetilde{w} = w_{n-1}^{(i)}\frac{p(z_n \mid \boldsymbol{x}_n)p(\boldsymbol{x}_n \mid \boldsymbol{x}_{n-1}^{(i)})}{q(\boldsymbol{x}_n \mid \boldsymbol{x}_{n-1}^{(i)},z_n)} \tag{16-55}$$

引入一个辅助变量 $\xi(\xi \in \{1,\cdots,N_s\}$,其索引是"先验"混合密度的一个分量。将联合条件索引后验密度定义为 $p(\boldsymbol{x}_n,\xi=i|z_{1:n})$,则有

$$\begin{aligned} p(\boldsymbol{x}_n,\xi = i \mid z_{1:n}) &\triangleq p(z_n \mid \boldsymbol{x}_n)p(\boldsymbol{x}_n,\xi = i \mid z_{1:n-1}) \\ &= p(z_n \mid \boldsymbol{x}_n)p(\boldsymbol{x}_n \mid \xi = i,z_{1:n-1})p(i \mid z_{1:n-1}) \end{aligned} \tag{16-56}$$

将上式与式(16-54)进行对比,得

$$p(\boldsymbol{x}_n \mid \xi = i,z_{1:n-1}) = p(\boldsymbol{x}_n \mid \boldsymbol{x}_{n-1}^{(i)}) \tag{16-57}$$

$$p(i \mid z_{1:n-1}) = w_{n-1}^{(i)} \tag{16-58}$$

已知 $\{z_n,\boldsymbol{x}_{n-1}^{(i)},w_{n-1}^{(i)}\}$,得 $\boldsymbol{x}_n$ 后验概率为

$$\beta_n^{(i)} \triangleq p(\boldsymbol{x}_n, \xi = i \mid z_{1:n}) = p(z_n \mid \boldsymbol{x}_n) p(\boldsymbol{x}_n \mid \boldsymbol{x}_{n-1}^{(i)}) w_{n-1}^{(i)} \tag{16-59}$$

通过忽略ξ，可以由联合密度(16－59)抽取样本。假设由边缘密度$p(\boldsymbol{x}_n \mid z_{1:n})$中抽取一组粒子$\{\boldsymbol{x}_n^{(i)}\}_{i=1}^{N_s}$，索引$\xi$用与$p(\xi \mid z_{1:n-1}) = w_{n-1}^{(i)}$成正比的概率模拟。这样，式(16－54)可由下式估计求得，即

$$p(\boldsymbol{x}_n \mid z_{1:n}) \propto \sum_{i=1}^{N_s} w_{n-1}^{(i)} p(z_n \mid \boldsymbol{x}_n, \xi^{(i)}) p(\boldsymbol{x}_n \mid \boldsymbol{x}_{n-1}^{(i)}) \tag{16-60}$$

式中：$\xi^{(i)}$为粒子$\boldsymbol{x}_n^{(i)}$在时间步t_{n-1}的索引(即$\xi^{(i)} \triangleq \{\xi = i\}$)。

用来抽取样本$\{\boldsymbol{x}_n^{(i)}, i = 1, \cdots, N_s\}$的建议分布，以一种分解的方式选取，即

$$q(\boldsymbol{x}_n, \xi \mid z_{1:n}) \propto q(\xi \mid z_{1:n}) q(\boldsymbol{x}_n \mid \xi, z_{1:n}) \tag{16-61}$$

$$q(\xi^{(i)} \mid z_{1:n}) = w_{n-1}^{(i)} p(z_n \mid \boldsymbol{\mu}_n^{(i)}) \tag{16-62}$$

$$q(\boldsymbol{x}_n \mid \xi^{(i)}, z_{1:n}) = p(\boldsymbol{x}_n \mid \boldsymbol{x}_{n-1}^{(i)}) \tag{16-63}$$

这里$\boldsymbol{\mu}_n^{(i)}$是与$p(\boldsymbol{x}_n \mid \boldsymbol{x}_{n-1}^{(i)})$相关的数值(例如均值、模式或样本值)，由此抽取第$i$个粒子。

由下式进一步逼近真正的后验概率，即

$$p(\boldsymbol{x}_n \mid z_{1:n}) \propto \sum_{i=1}^{N_s} w_{n-1}^{(i)} p(z_n \mid \boldsymbol{\mu}_n^{(\xi=i)}) p(\boldsymbol{x}_n \mid \boldsymbol{x}_{n-1}^{(\xi=i)}) \tag{16-64}$$

这样式(16－55)的非规范化权值变成

$$\tilde{w}_n^{(i)} = w_{n-1}^{(\xi=i)} \frac{p(z_n \mid \boldsymbol{x}_n^{(i)}) p(\boldsymbol{x}_n^{(i)} \mid \boldsymbol{x}_{n-1}^{(\xi=i)})}{q(\boldsymbol{x}_n^{(i)}, \xi^{(i)} \mid z_{1:n})} = \frac{p(z_n \mid \boldsymbol{x}_n^{(i)})}{p(z_n \mid \boldsymbol{\mu}_n^{(\xi=i)})} \tag{16-65}$$

表16－6给出了辅助粒子滤波器的过程流程表。为总结SIS APF的处理流程，第2步相当于使用自举粒子滤波器计算一组新的权值，然后使用这些新的权值对粒子集$\{\boldsymbol{x}_{n-1}^{(i)}, i = 1, \cdots, N_s\}$进行重采样和正则化。继而在第5步重采样后的粒子进行时间向前传播，利用式(16－65)计算一组规范化的新权值。

表16－6　辅助粒子滤波器过程流程表

1. 滤波器初始化 (1) 初始化状态矢量样本； (2) 初始化权值	$\boldsymbol{x}_0^{(i)} \sim q(\boldsymbol{x}_0) \quad i = 1, \cdots, N_s$ $w_0^{(i)} = \frac{1}{N_s}$
2. 计算	$\mu_n^{(i)}$ mean or draw from $p(\boldsymbol{x}_n \mid \boldsymbol{x}_{n-1}^{(i)}), \forall i$ $\tilde{\beta}_n^{(i)} = w_{n-1}^{(i)} p(z_n \mid \boldsymbol{\mu}_n^{(i)}), \forall i$
3. 规范化权值	$\beta_n^{(i)} = \tilde{\beta}_n^{(i)} / \sum_{i=1n}^{N_s} \tilde{\beta}_n^{(i)}$

（续）

4. 利用$\beta_n^{(i)}$重采样$\boldsymbol{x}_{n-1}^{(i)}$ 移动重采样粒子	利用 MATLAB 重采样例程进行重采样$\boldsymbol{x}_{n-1}^{(i)}$用$\beta_n^{(i)}$取代$w_{n-1}^{(i)}$ 用表15－3中的步骤对粒子进行正则化
5. 重要性采样	$\boldsymbol{x}_n^{(i)} = p(\boldsymbol{x}_n \mid \boldsymbol{x}_{n-1}^{(i)}, \xi^{(i)})$ $w_n^{(i)} = \dfrac{p(z_n \mid \boldsymbol{x}_n^{(i)})}{p(z_n \mid \boldsymbol{\mu}_n^{(\xi=i)})}$
6. 规范化	$w_n^{(j)} = \widetilde{w}_n^{(j)} / \sum_{i=1}^{N_s} \widetilde{w}_n^{(j)}$
7. 时间递增，返回第2步	$\boldsymbol{x}_n^{(j)} \rightarrow \boldsymbol{x}_{n-1}^{(i)}; w_n^{(j)} \rightarrow w_{n-1}^{(i)}$

APF 基本上分为两个阶段：在第一阶段，仿真具有大的预测似然性的粒子；在第二阶段，重新计算粒子权值，抽取增大状态，这相当于做一个建议，选高条件似然性的粒子作为先验，以此避免无效的抽样。因此，APF 利用从似然模型中得的事前信息，以避免无效抽样，因为低似然性的粒子中包含的信息很少。换句话说，就是直接将粒子抽样推广到高似然度区域。

以下为关于 APF 的一些论述：

（1）在一般 SIS 粒子滤波器中，通常是在重采样步骤之后进行估计，这样效率很低，因为重采样引入了额外的当前状态的随机变化。APF 通过一步基于点估计$\boldsymbol{\mu}_n^{(i)}$表示$p(\boldsymbol{x}_n|\boldsymbol{x}_n^{(i)})$的预估计克服了这一问题。

（2）当过程噪声很小时，APF 的性能通常优于 SIS 滤波器。然而，当过程噪声很大时，点估计$\boldsymbol{\mu}_n^{(i)}$无法提供足够的$p(\boldsymbol{x}_n|\boldsymbol{x}_n^{(i)})$信息，这样就无法保证 APF 的优势。

（3）在 APF 中，重要性密度建议为混合密度，依赖于过去的状态和最近的观测值。

（4）APF 的思想与文献[18]中提出的局部蒙特卡洛方法相同，作者给出两种方法实现基于联合分布或边缘分布的$\{\boldsymbol{x},\xi\}$样本抽样。

（5）APF 的缺点是从增大的（因此更高的）空间中进行样本抽样，如果辅助索引与固定先验相比差异很大，这种增大可以忽略，重要性权值的方差会更高。

16.5.1 APF 在 DIFAR 浮标跟踪中的应用

在 DIFAR 实例中，采用 APF 作为跟踪滤波器，希望跟踪结果与高斯卡尔曼滤波器相比能有所改善。但如图 16－6 所示，对 6 个信噪比递减的输入信号进行蒙特卡洛跟踪结果进行对比，APF 并没有表现出任何优势。造成这种现象的原因是显而易见的。粒子滤波器是在当后验密度（表示动态模型）为非高斯时

实现估计。但是对于 DIFAR 实例,动态模型为线性高斯而似然函数是非高斯的。因此,在这种情况下,APF 并不优于高斯卡尔曼滤波器。

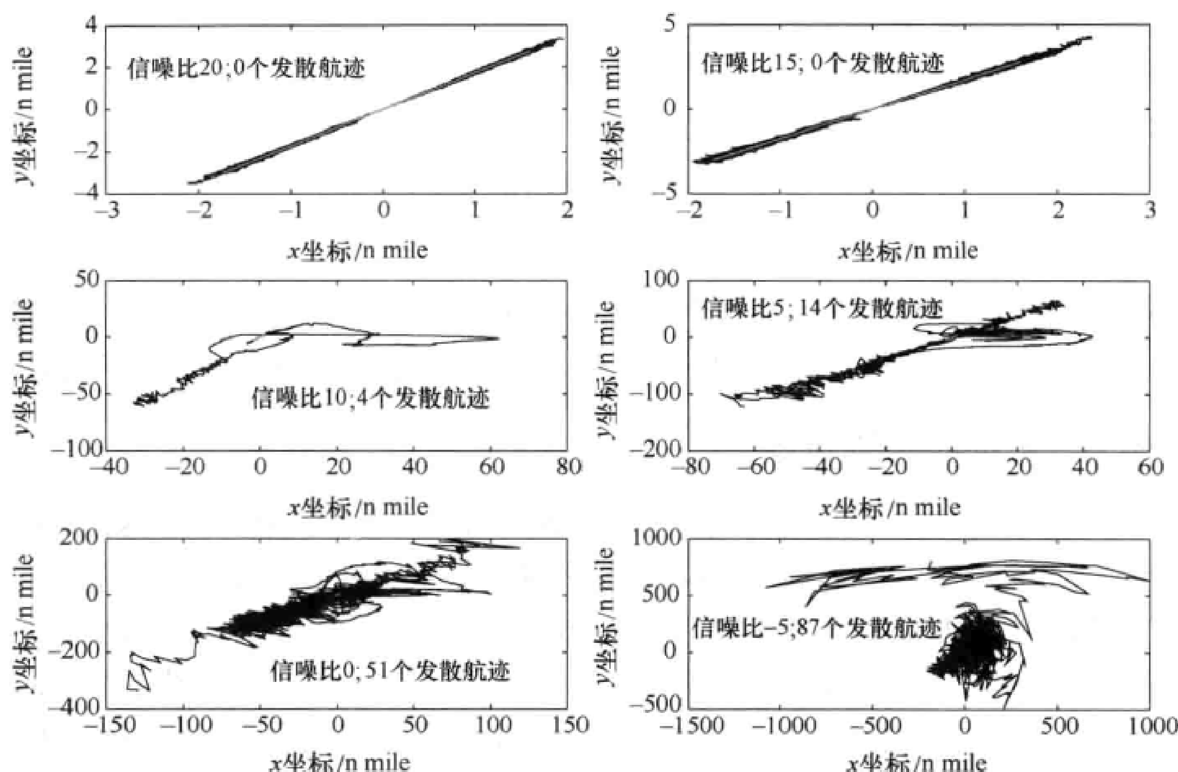


图 16-6　将辅助粒子滤波器应用于 DIFAR 跟踪实例的跟踪估计结果

16.6　序贯重要性采样辅助粒子滤波器估计

下面给出另一种 SIS 粒子滤波方法,与上面给出的 APF 非常类似。假设用高斯密度估计得后验密度 $p(\boldsymbol{x}_n|\boldsymbol{z}_{1:n})$ 和最优重要性密度 $q(\boldsymbol{x}_n^{(i)}|\boldsymbol{x}_{n-1}^{(i)},\boldsymbol{z}_{1:n})$,后验密度估计为

$$p(\boldsymbol{x}_n \mid \boldsymbol{z}_{1:n}) \approx \mathcal{N}(\boldsymbol{x}_n;\bar{\boldsymbol{x}}_n,\hat{\boldsymbol{P}}_n) \tag{16-66}$$

用高斯建议分布来近似每个粒子的建议分布,即

$$q(\boldsymbol{x}_n^{(i)} \mid \boldsymbol{x}_{n-1}^{(i)},\boldsymbol{z}_{1:n}) \approx \mathcal{N}(\boldsymbol{x}_n^{(i)};\bar{\boldsymbol{x}}_n^{(i)},\hat{\boldsymbol{P}}_n^{(i)}) \tag{16-67}$$

即在 t_{n-1} 时刻,利用本书第Ⅱ部分中给出的高斯卡尔曼滤波器的变体,根据新的观测值,计算每个粒子重要性分布的均值和协方差。然后,在这个分布中进行第 i 个粒子的采样。

16.6.1　扩展卡尔曼粒子滤波器

在 Merwe 等[19]首次提出的扩展卡尔曼粒子滤波器中,将非线性动态方程和观测方程展开为泰勒级数,并应用于 EKF 当中,以生成第 i 个粒子的前两阶矩

$(\bar{\boldsymbol{x}}_n^{(i)}, \hat{\boldsymbol{P}}_n^{(i)})$,将其用在式(16-67)中来定义高斯建议分布。然后,从建议分布中抽取第 i 个样本,即

$$\hat{\boldsymbol{x}}_n^{(i)} = q(\boldsymbol{x}_n^{(i)} \mid \boldsymbol{x}_{n-1}^{(i)}, \boldsymbol{z}_{1:n}) = \mathcal{N}(\boldsymbol{x}_n^{(i)}; \bar{\boldsymbol{x}}_n^{(i)}, \hat{\boldsymbol{P}}_n^{(i)}) \tag{16-68}$$

将 $\hat{\boldsymbol{x}}_n^{(i)}$ 代入式(16-12)来更新权重,即

$$\tilde{w}_n^{(i)} = w_{n-1}^{(i)} \frac{p(\boldsymbol{z}_n \mid \hat{\boldsymbol{x}}_n^{(i)}) p(\hat{\boldsymbol{x}}_n^{(i)} \mid \boldsymbol{x}_{n-1}^{(i)})}{q(\hat{\boldsymbol{x}}_n^{(i)} \mid \boldsymbol{x}_{n-1}^{(i)}, \boldsymbol{z}_n)} \tag{16-69}$$

这样就得了规范化权值。因为这个方法与前面提出的高斯最优 SIS 粒子滤波器几乎完全相同,读者可自行参阅文献[19],以获得此方法相关具体步骤的详细信息。

16.6.2 无迹粒子滤波器

文献[19]提出了无迹粒子滤波器(UPF),对于高度非线性系统 UKF 已被证实比卡尔曼滤波器更为精确。用 UKF 取代 EKF,生成用于高斯建议分布的前两阶矩。UPF 方法非常简单、容易实现,表 16-7 给出了详细步骤,可以用表 13-2 来计算 Sigma 点权重。此外,表中省略了一些与前面的粒子滤波器完全相同的步骤。UPF 相关信息可参阅文献[20,21],UPF 的平方根变体可参阅文献[22]。

表 16-7 带重采样和正则化的无迹粒子滤波器

1. 滤波器初始化 (1) 初始化状态矢量样本; (2) 初始化权值	$\boldsymbol{x}_0^{(i)} \sim q(\boldsymbol{x}_0) \quad i = 1, \cdots, N_s$ $w_0^{(i)} = \frac{1}{N_s}$
2. 重要性采样 For $i = 1, \cdots, N_s$ (1) 计算第 j 个 Sigma 点; (2) 状态预测	选取 UKF 权值 w_0 $\boldsymbol{\chi}_{n-1\mid n-1}^{(i,j)} = \bar{\boldsymbol{x}}_{n-1\mid n-1}^{(i)} + \sqrt{\frac{n_x}{1-w_0}\boldsymbol{P}_{n-1\mid n-1}^{xx(i)}}\boldsymbol{r}^{(j)} \quad j = 0, 1, \cdots, 2n_x$ $\hat{\boldsymbol{x}}_{n\mid n-1}^{(i)} = \sum_{j=0}^{2n_x} w_j \boldsymbol{f}(\boldsymbol{\chi}_{n-1\mid n-1}^{(i,j)})$ $\boldsymbol{P}_{n\mid n-1}^{xx(i)} = \sum_{j=0}^{2n_x} w_j \boldsymbol{f}(\boldsymbol{\chi}_{n-1\mid n-1}^{(i,j)}) \boldsymbol{f}^{\mathrm{T}}(\boldsymbol{\chi}_{n-1\mid n-1}^{(i,j)}) - \hat{\boldsymbol{x}}_{n\mid n-1}^{(i)} \hat{\boldsymbol{x}}_{n\mid n-1}^{\mathrm{T}(i)} + \boldsymbol{Q}$
3. 观测值预测	$\boldsymbol{\chi}_{n\mid n-1}^{(i,j)} = \hat{\boldsymbol{x}}_{n\mid n-1}^{(i)} + \sqrt{\frac{n_x}{1-w_0}\boldsymbol{P}_{n\mid n-1}^{xx(i)}}\boldsymbol{r}^{(j)} \quad j = 0, 1, \cdots, 2n_x$ $\hat{\boldsymbol{z}}_{n\mid n-1}^{(i)} = \sum_{j=0}^{2n_x} w_j \boldsymbol{h}(\boldsymbol{\chi}_{n\mid n-1}^{(i,j)})$ $\boldsymbol{P}_{n\mid n-1}^{zz(i)} = \sum_{j=0}^{2n_x} w_j \boldsymbol{h}(\boldsymbol{\chi}_{n-1\mid n-1}^{(i,j)}) \boldsymbol{h}^{\mathrm{T}}(\boldsymbol{\chi}_{n-1\mid n-1}^{(i,j)}) - \hat{\boldsymbol{z}}_{n\mid n-1}^{(i)} \hat{\boldsymbol{z}}_{n\mid n-1}^{\mathrm{T}(i)} + \boldsymbol{R}$ $\boldsymbol{P}_{n\mid n-1}^{xz(i)} = \sum_{j=0}^{2n_x} w_j \boldsymbol{f}(\boldsymbol{\chi}_{n-1\mid n-1}^{(i,j)}) \boldsymbol{h}^{\mathrm{T}}(\boldsymbol{\chi}_{n-1\mid n-1}^{(i,j)}) - \hat{\boldsymbol{x}}_{n\mid n-1}^{(i)} \hat{\boldsymbol{z}}_{n\mid n-1}^{(i)}$

（续）

4. 更新卡尔曼滤波器	$K_n^{(i)} = P_{n\mid n-1}^{xz(i)}[P_{n\mid n-1}^{zz(i)}]^{-1}$ $\hat{x}_{n\mid n}^{(i)} = \hat{x}_{n\mid n-1}^{(i)} + K_n^{(i)}(z_n - \hat{z}_{n\mid n-1}^{(i)})$ $P_{n\mid n}^{xx(i)} = P_{n\mid n-1}^{xx(i)} - K_n^{(i)} P_{n\mid n-1}^{zz(i)} K_n^{\mathrm{T}(i)}$
5. 更新重要性权值	$\tilde{w}_n^{(i)} = w_{n-1}^{(i)} \dfrac{p(z_n \mid \hat{x}_n^{(i)})p(\hat{x}_n^{(i)} \mid x_{n-1}^{(i)})}{q(\hat{x}_n^{(i)} \mid x_{n-1}^{(i)}, z_n)}$ $w_n^{(j)} = \tilde{w}_n^{(j)} / \sum_{i=1}^{N_s} \tilde{w}_n^{(j)}$
6. 重采样并移动 7. 矩计算 8. 时间递增，返回第 2 步	

注意：UKF 只是可用于此类应用的 Sigma 点卡尔曼滤波器类中的一种。可以很容易地提出一种球面单形粒子滤波器或者高斯－厄米特 SIS 粒子滤波器，将这些滤波器的权值和 Sigma 点代入表 16－7 的运算步骤中。实际上，对于高维状态矢量，用球面单形卡尔曼滤波器来代替无迹卡尔曼滤波器是一个好思路。另一方面，对于低维状态矢量和高度非线性模型，换成高斯－厄密特卡尔曼滤波器会更有意义。为保证完整性，甚至可以在这种应用中使用蒙特卡洛卡尔曼滤波器，这样不会损失任何一般性。

16.7 利用 Rao－Blackwellization 理论降低计算负荷

在许多状态估计的应用中，状态矢量是高维的，SIS 粒子滤波器的计算负荷变得极大，估计精度可能迅速降低。当应用于高维状态空间时，因为需要大数量的样本来保证密度函数的估计的精确性，基于粒子滤波的方法的效率变得极低。状态矢量的分量可以细分为两类：一类遵循线性时域迁移，是高斯的；第二类是非高斯的，可利用 Rao－Blackwellization[23－25] 来减少计算负荷。

Rao－Blackwell 算法的关键思想是降维，利用模型结构将条件后验分成两个部分，其中高斯部分以非高斯部分为条件。首先用粒子滤波器生成非高斯后验，然后对条件高斯部分应用卡尔曼滤波器。假设状态矢量分量可以被分成两组 $x_n = [q^{\mathrm{T}}, r^{\mathrm{T}}]^{\mathrm{T}}$，这样有

$$p(x_n \mid z_{1:n}) = p(q_n, r_n \mid z_{1:n}) = p(q_n \mid r_n, z_{1:n})p(r_n \mid z_{1:n}) \qquad (16-70)$$

在文献[24]中指出，当 q_n 的初始不确定性为高斯分布，而 q_n 的观测模型和系统动态的条件概率也为高斯分布时，可以用线性卡尔曼滤波器有效地更新 $p(q_n \mid r_n, z_{1:n})$。为深入理解并使用 Rao－Blackwell 方法，读者可自行参阅文献[3－24－25]。

参考文献

[1] Riple BD. *Stochastic Simulation*. Wiley,1987.

[2] Online at: http://en.wikipedia.org/wiki/Inverse.transform_sampling.

[3] Ristic B, Arulampalam S, Gordon N. Beyond the Kalman Filter: Particle Filters for Tracking Applications. Artech House,2004.

[4] Andrieu C, Douct A, Holstein R. Particle Markov chain Monte Carlo methods. *J. R. Stat. Soc. B*, 2010, 72(3):269 - 324.

[5] Cappe O, Godsill SJ, Moulines E. An overview of existing methods and recent advances in sequential Monte Carlo. *Proc. IEEE*,2007,95(5):899 - 924.

[6] Liu JS, Chen R. Sequential Monte Carlo methods for dynamic systems. *J. Am. Stat. Assoc*, 1998,93 (443): 1032 - 1044.

[7] Chen Z. Bayesian Filtering: From Kalman Filters to Particle Filters, and Beyond. Adaptive Systems Laboratory, McMaster University, Hamilton, ON, Canada. [Online], http://citeseerx.ist.psu.edu.

[8] Fearnhead P. Sequential Monte Carlo Methods in Filter Theory. Dissertation. Oxford UK: University of Oxford. MertonCollege,1998.

[9] Ross SM. *Introduction to Probability Models*. 4th ed. Academic Press, 1989.

[10] Hoi DH, Schon TB. Gustafsson F. On resampling algorithms for particle filters. In *Proceedings of Nonlinear Statistical Signal Processing Workshop*, Cambridge, UK, 2006,9.

[11] Kong A, Liu JS, Wong WH. Sequential computations and Bayesian missing data problems. *J. Am. Stat. Assoc*, 1994,89(425):278 - 288.

[12] Gordon N, Salmond D, Smith AF. Novel approach to nonlinear/non - Gaussian Bayesian state estimation. *IEE Proc. F. Radar Signal Process*, 1993, 140:107 - 113.

[13] Doucet A. Monte Carlo Methods for Bayesian Estimation of Hidden Markov Models. Application to Radiation Signals, Dissertation. University of Paris - Sud, Orsay, 1997.

[14] Doucet A. On Sequential Simulation - Based methods for Bayesian Filtering. Technical Report CUED/F - INFENG/TR.310, University of Cambridge,1998.

[15] Doucet A, Gordon N, Krishnamurthy V. Particle filters for state estimation of jump Markov linear systems. *IEEE Trans. Sig. Proc*, 2001,49:613 - 624.

[16] Pitt K, Shephard N. Filtering via simulation: auxiliary particle filter. *J. Am. Stat. Assoc*,1999,94(446): 590 - 599.

[17] Oitt MK, Shephard N. Auxiliary variable based particle filters. In *Sequential Monte Carlo Methods in Practice*. Springer, 2001.

[18] Wu W. Hu X. Hu D. Wu M. Comments on "Gaussian Particle Filtering". *IEEE Trans. Sig. Proc*, 2005, 53(8):3350 - 3351.

[19] van der Merwe R. Doucet A, de Freitas N, Warn E. The Unscented Particle Filter. Cambridge University Technical Report CUED/F - INFENG/TR 380,2000.

[20] van der Merwe R, Doucet A, de Freitas N, et al. The unscented particle filter. *Adv. Neural Inform. Process. Syst*, 2001,13:584 - 590.

[21] Rui Y, Chen Y. Better Proposal distributions: object tracking using unscented particle filter. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, CVPR 2001, 2001(2):786 - 793.

[22] Zandara S, Nicholson AE. Square Root Unscented Particle Filtering for Grid Mapping. 6th *World Congress on Intelligent Control and Automation*, 2006:1548 - 1552.

[23] Bolviken E, Storik G. Deterministic and Stochastic Particle Filters for State - Space Models. In *Sequential Monte Carlo Methods in Practice*. Springer, 2001.

[24] Murphy K, Russell S. Rao - Blackwellised particle filtering for dynamic Bayesian networks. In *Sequential Monte Carlo Methods in Practice*. Springer, 2001.

[25] Casella C, Rober CP. Rao - Blackwellisation of sampling schemes. *Biometrica*, 1996, 83(1):81 - 94.

第17章 广义蒙特卡洛粒子滤波器

所有SIS类粒子滤波器都是基于式(16-12),该式定义了重要性权值的递归更新。将递归重要性权值更新与重采样和移动步骤相结合,提供了一个非线性非高斯系统的估计方法框架,只需要选择重要性密度和似然函数。但是,SIS粒子滤波器在使用中有许多缺点:

(1) BPF在重要性采样过程中不使用最新的测量值,使得它容易受许多应用中方差和不稳定性增加的影响。

(2) 所有SIS粒子滤波器都需要重采样和移动步骤,来增加估计状态误差协方差。SIS粒子滤波器不能并行化,因此还需要增加一个重要的计算负荷估计步骤。

(3) SIS APF及其替代算法可能需要对每个粒子在每个时间步执行高斯卡尔曼滤波器,进一步提高了计算负荷。

第二类粒子滤波器是广义蒙特卡洛(GMC)粒子滤波器,它不使用递归的权值更新,而是在每次序贯更新滤波器时计算新的权值。这类更为通用的粒子滤波器的实例包括Kotechha和Djuric[1,2]提出的高斯粒子滤波器。

17.1 高斯粒子滤波器

再次回顾贝叶斯估计的基本方程,记住这是一个两步过程,滤波步骤由后验估计(滤波或更新)密度得出,即

$$p(\boldsymbol{x}_n \mid \boldsymbol{z}_{1:n}) = cp(\boldsymbol{z}_n \mid \boldsymbol{x}_n)p(\boldsymbol{x}_n \mid \boldsymbol{z}_{1:n-1}) \tag{17-1}$$

预测步骤包括预测密度(或先验密度,因为它使用之前的所有测量值)的估计,即

$$p(\boldsymbol{x}_n \mid \boldsymbol{z}_{1:n-1}) = \int p(\boldsymbol{x}_n \mid \boldsymbol{x}_{n-1})p(\boldsymbol{x}_{n-1} \mid \boldsymbol{z}_{1:n-1})\mathrm{d}\boldsymbol{x}_{n-1} \tag{17-2}$$

GPF 的基本概念是假设后验密度和先验密度都是高斯分布,即

$$p(\boldsymbol{x}_n \mid z_{1:n}) = \mathcal{N}(\boldsymbol{x}_n;\boldsymbol{\mu}_n,\boldsymbol{\Sigma}_n) \tag{17-3}$$

$$p(\boldsymbol{x}_n \mid z_{1:n-1}) = \mathcal{N}(\boldsymbol{x}_n;\bar{\boldsymbol{\mu}}_n,\bar{\boldsymbol{\Sigma}}_n) \tag{17-4}$$

假设在 t_0 时刻,在所有观测值之前,有如下初始密度的先验信息,即

$$p(\boldsymbol{x}_0) = \mathcal{N}(\boldsymbol{x}_0;\boldsymbol{\mu}_0,\boldsymbol{\Sigma}_0) \tag{17-5}$$

为初始化 GPF,首先由 $p(\boldsymbol{x}_0)$ 抽取样本 $\{\tilde{\boldsymbol{x}}_0^{(i)}\}_{i=1}^{N_s}$,然后使用动态方程将样本在时间上向前传播,即

$$\tilde{\boldsymbol{x}}_{n|n-1}^{(i)} = \boldsymbol{f}_n(\tilde{\boldsymbol{x}}_{n-1|n-1}^{(i)}) + \boldsymbol{\mu}_n \tag{17-6}$$

然后计算 $\bar{\boldsymbol{\mu}}_1$ 和 $\bar{\boldsymbol{\Sigma}}_1$,即

$$\bar{\boldsymbol{\mu}}_n = \frac{1}{N_s}\sum_{i=1}^{N_s}\tilde{\boldsymbol{x}}_{n|n-1}^{(i)} \tag{17-7}$$

$$\bar{\boldsymbol{\Sigma}}_n = \frac{1}{N_s}\sum_{i=1|n|n-1}^{N_s}\tilde{\boldsymbol{x}}_{n|n-1}^{(i)}\tilde{\boldsymbol{x}}_{n|n-1}^{(i)\mathrm{T}} - \bar{\boldsymbol{\mu}}_n\bar{\boldsymbol{\mu}}_n^{\mathrm{T}} + \boldsymbol{Q} \tag{17-8}$$

这定义了式(17-4)的先验(预测)密度。

回顾式(15-41),可以将重要性权值计算方法写成非迭代的形式,即

$$w_n \propto \frac{p(z_n \mid \boldsymbol{x}_n)\mathcal{N}(\boldsymbol{x}_n;\bar{\boldsymbol{\mu}}_n,\bar{\boldsymbol{\Sigma}}_n)}{q(\boldsymbol{x}_n \mid z_{1:n})} \tag{17-9}$$

在从(当前未指明的)重要性密度中采样后,$\boldsymbol{x}_n^{(i)} \sim q(\boldsymbol{x}_n \mid z_{1:n}),i=1,\cdots,N_s$,权值变成

$$\tilde{w}_n^{(i)} = \frac{p(z_n \mid \boldsymbol{x}_n = \boldsymbol{x}_n^{(i)})\mathcal{N}(\boldsymbol{x}_n = \boldsymbol{x}_n^{(i)};\bar{\boldsymbol{\mu}}_n,\bar{\boldsymbol{\Sigma}}_n)}{q(\boldsymbol{x}_n = \boldsymbol{x}_n^{(i)} \mid z_{1:n})} \tag{17-10}$$

经过规范化步骤,得

$$w_n^{(i)} = \frac{\tilde{w}_n^{(i)}}{\sum_{i=1}^{N_s}\tilde{w}_n^{(i)}} \tag{17-11}$$

可以计算后验密度的矩,即

$$\hat{\boldsymbol{x}}_n = \sum_{i=1}^{N_s} w_n^{(i)}\boldsymbol{x}_n^{(i)} \tag{17-12}$$

$$\boldsymbol{P}_n^{xx} = \sum_{i=1}^{N_s} w_n^{(i)}\boldsymbol{x}_n^{(i)}\boldsymbol{x}_n^{(i)\mathrm{T}} - \hat{\boldsymbol{x}}_n\hat{\boldsymbol{x}}_n^{\mathrm{T}} \tag{17-13}$$

这样就完成了高斯粒子滤波器过程,如图 17-1 所示。注意,广义蒙特卡洛 GPF 仍然需要额外的似然函数和重要性密度的解析表达式。

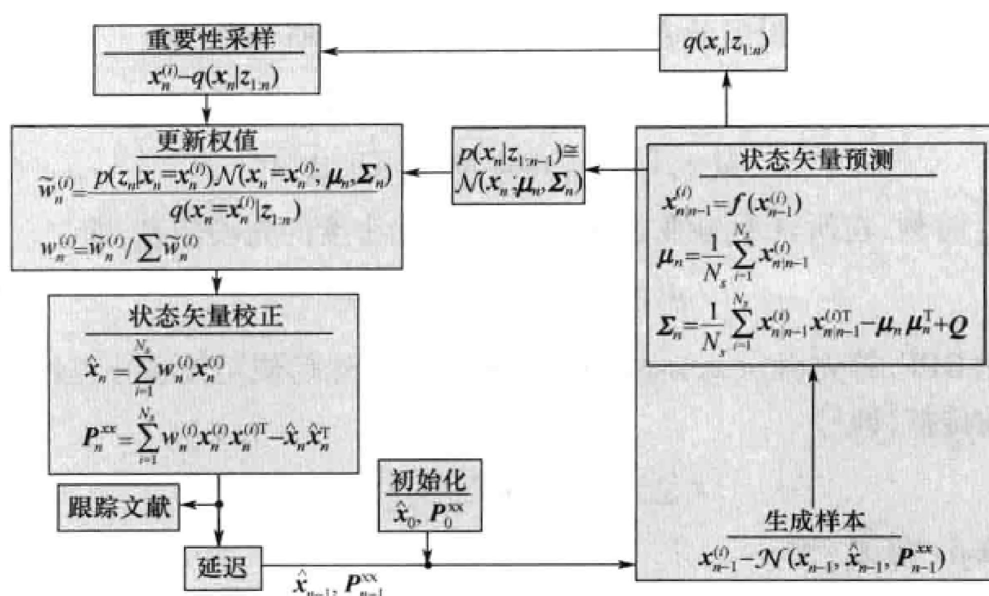


图 17-1　高斯粒子滤波器过程流图

17.2　组合粒子滤波器

图 17-1 并未指明重要性密度。因此,GPF 是不完整的,需要指定一个重要性密度。在文献[1] 中,建议采用高斯分布作为重要性密度,即令 $q(\boldsymbol{x}_n \mid \boldsymbol{z}_{1:n}) = \mathcal{N}(\boldsymbol{x}_n;\boldsymbol{\mu}_{n|n}^x,\Sigma_{n|n}^{xx})$,其中$\boldsymbol{\mu}_{n|n}^x$ 和 $\Sigma_{n|n}^{xx}$ 用本书第 Ⅱ 部分的任意一种非线性卡尔曼滤波器,由先验密度求解。这种方法将 GPF 与一种非线性卡尔曼滤波器结合起来,因此称为广义蒙特卡洛组合粒子滤波器(CPF)。

可以将 CPF 划分为两类:利用 EKF 来生成重要性密度和利用 Sigma 点卡尔曼滤波器。Sigma 点卡尔曼滤波器类可以进一步分成 4 种数值积分卡尔曼滤波器:MCKF、UKF、SSKF 或 GHKF。图 17-2 给出了 CPF-EKF 估计滤波器递归过程流图,图 17-3 给出了利用 Sigma 点卡尔曼滤波器的 CPF 过程流图。在图的下端可看到高斯粒子滤波器的结构,而在图的右侧,可以看到非线性卡尔曼滤波器结构。在参考文献[3] 中,提出一种类似的无迹粒子滤波器,但是不包括先验的高斯估计。需要注意文献[3] 包含大量的错误,使它非常难以遵循。

广义蒙特卡洛组合粒子滤波器在多数应用中运行良好,这是因为与重要性密度相关的方差通常比较宽,包括后验密度支撑的主要部分。除了后验为重尾分布的应用,对于任何应用来说,组合粒子滤波器是近似“最优” 的粒子滤波器。因为组合粒子滤波器不需要采样和移动步骤,而且每次迭代仅执行一次非线性卡尔曼滤波器,计算效率显著提高,优于大部分的 SIS 粒子滤波器。

对先验、后验密度重尾或多峰情况,有多种备选方案,可以用来提高广义蒙特卡洛粒子滤波器类的性能。针对多峰非高斯密度,Kotecha 和 Djuric[4] 提出一种高斯混合粒子滤波器,其中预测和滤波分布均近似高斯混合。在许多应用中,

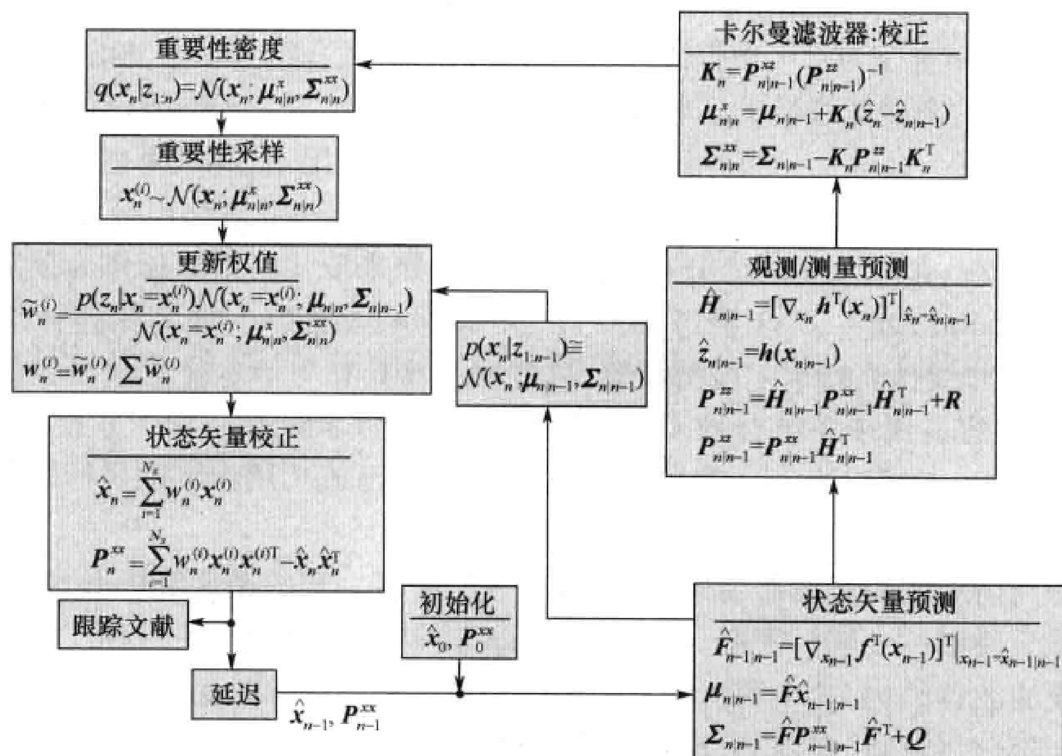


图 17 - 2　利用 EKF 作为重要性密度的组合粒子滤波器

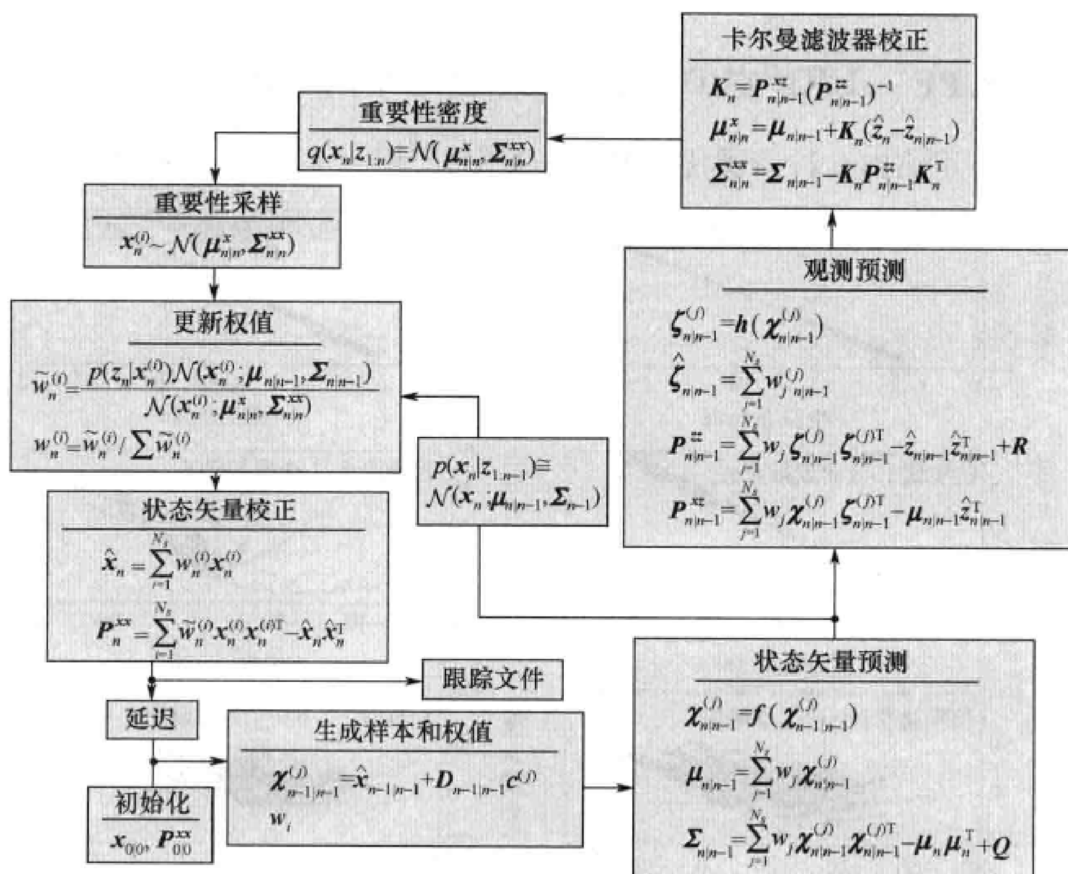


图 17 - 3　利用 Sigma 点卡尔曼滤波器作为重要性密度的组合粒子滤波器

这些高斯求和粒子滤波器(Gaussian Sum Particle Filters,GSPF)提高了性能。Merwe和Wan在文献[5]中提出一种Sigma点高斯混合组合粒子滤波器,根据受γ噪声影响的高度非线性动态转移方程估计一维状态。带γ动态噪声的状态矢量得增强,即便在重尾γ分布导致很多异常观测值的情况下,其结果也非常可靠。

可以用一些广义蒙特卡洛方法的改进方法,来处理重尾噪声分布。例如:使用Levy分布、alpha - stable分布或拉普拉斯分布来取代高斯逼近作为预测和滤波密度;用Kalman - Levy滤波器[6]、广义stable滤波器[7]或者拉普拉斯滤波器[8]取代非线性高斯卡尔曼滤波器。这些滤波器具有与卡尔曼滤波器类似的对称结构,很容易用来在组合粒子滤波器应用中代替非线性高斯卡尔曼滤波器。Kalman - Levy滤波器[9,10]和拉普拉斯滤波器[8]已成功用作机动目标跟踪滤波器,但尚未被用在粒子滤波器应用中。

和高斯卡尔曼滤波器类似,所有这些方法需要一致的密度解析形式。所有的滤波器只使用一个解析密度。Chen和Singpurwalla[11]提出了另一种广义非线性卡尔曼滤波器结构,其中解析密度会根据滤波器的不同部分而变化。对于滤波器的不同部分,包括状态预测、观测值预测和卡尔曼滤波器更新形式,利用γ分布、β分布和Pearson分布,始终坚持更新卡尔曼滤波器的公式表示。在实际的应用中,使用第12章中给出的一种与MCKF相似的滤波器,其方法可以很容易应用在广义粒子滤波器形式上。

17.2.1 CPF - UKF在DIFAR浮标跟踪中的应用

图17 - 4给出6种信噪比下CPF - UKF应用于DIFAR跟踪问题得到的蒙

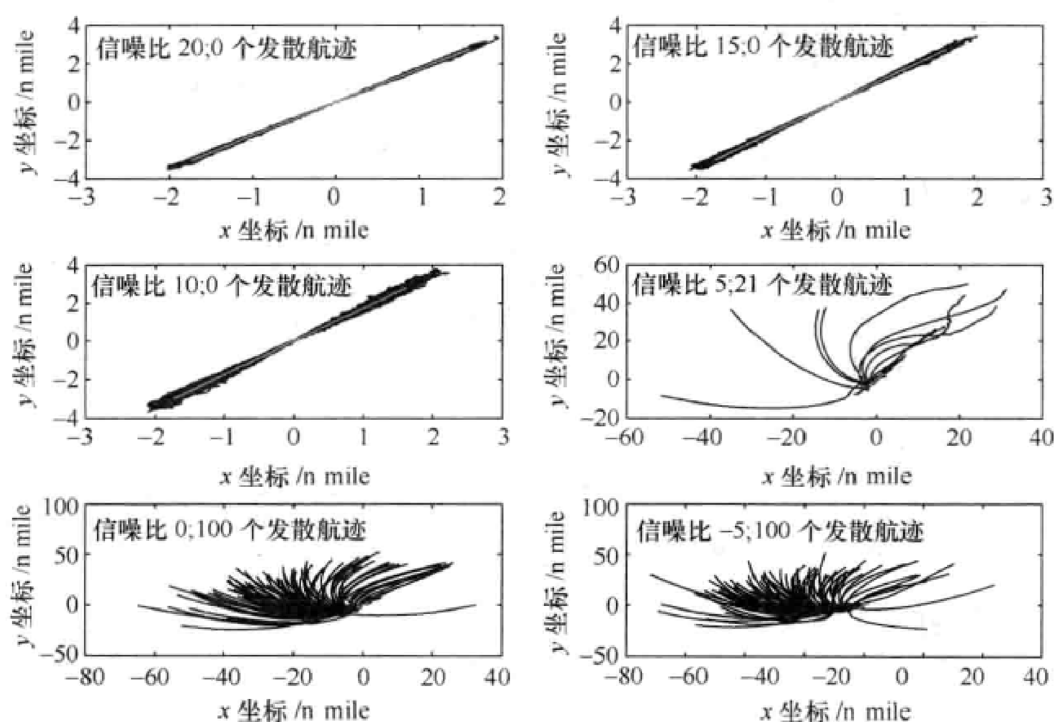


图17 - 4　6种信噪比下Sigma点高斯滤波器航迹图

特卡洛跟踪结果。与之前其它跟踪滤波器进行比较,可以看出这种滤波器并不优于(也不劣于)其它任何滤波器。

17.3 DIFAR 跟踪滤波器的性能对比

本节将一些跟踪滤波器应用于 DIFAR 实例,并分析其性能。本书把这种性能分析放到在第 Ⅱ 和第 Ⅲ 部分介绍完所有的滤波器之后。因为到目前为止,所有的结果都是基于仿真观测数据,而最好的性能估计方法是比较不同滤波器的 x 坐标位置和 y 坐标位置的 RMSE。对于每种信号信噪比,给出 4 种高斯滤波器(EKF、UKF、SSKF 和 GHKF)和 3 种粒子滤波器(BPF、APF 和 GSPF)的 RMS 位置误差比较图。GSPF 将粒子滤波器与 UKF 组合作为重要性密度。对于所有的粒子滤波器,使用 3000 个蒙特卡洛样本。

作为第一个实例,图 17-5 给出信号信噪比为 20dB 时,将 x 轴和 y 轴均方根误差作为目标船舶的真实位置函数的比较。可以看到,与图 16-5 一样,在信噪比为 20dB 时,由于有 3 个发散航迹(从 100 次蒙特卡洛运行中得),BPF 的 RMS 位置误差迅速增加。这只是关于 BPF 初值敏感性的另一个例证。因为所有的滤波器都采用同一组初值进行初始化,BPF 所需的这种额外的调整对于真实世界跟踪应用来说,是不可能实现的。

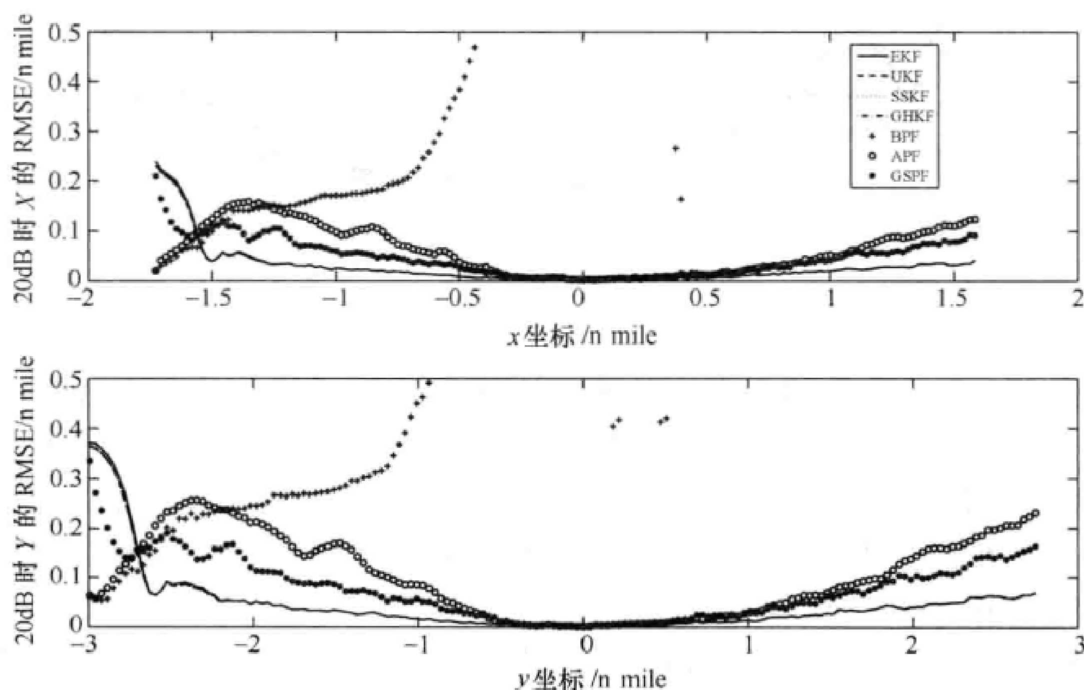


图 17-5 当输入信号信噪比为 20dB 时,DIFAR 实例均方根位置误差比较

注意,所有的高斯卡尔曼滤波器变体的性能都几乎相同,略优于 APF 和 GSPF。前面章节已暗示,在所有粒子滤波器中,GSPF 的性能最好。图中还可以

看出，大多数滤波器在跟踪刚刚开始时 RMSE 跟踪性能指标最高，当船舶驶入浮标场时达到最小值，几乎是关于原点对称的。最初的高 RMSE 是由于滤波器初始值选取得不好，但当滤波器综合一些观测值之后，就可以回归到合理的跟踪估计。

如图 17－6 所示，当信号信噪比降至 15dB 时，相对于 20dB 的情况，RMS 位置误差增加。由于在这个信噪比下有 7 个发散航迹，所以 BPF 的位置 RMSE 是发散的。在 100 次蒙特卡洛运行中，其它滤波器都没有任何发散航迹。在高信噪比下，高斯滤波器的 RMS 位置误差与 APF 和 GSPF 相比几乎相同，稍高一点。图 14－3 至图 14－7 给出信号信噪比由 20dB 降至 0dB 时高斯滤波器的 RMS 位置误差。这些图中的高斯滤波器的 RMS 误差均按同一比例，因此可以更好地比较高斯滤波器误差，也能表明高斯滤波器之间细微的性能差异。

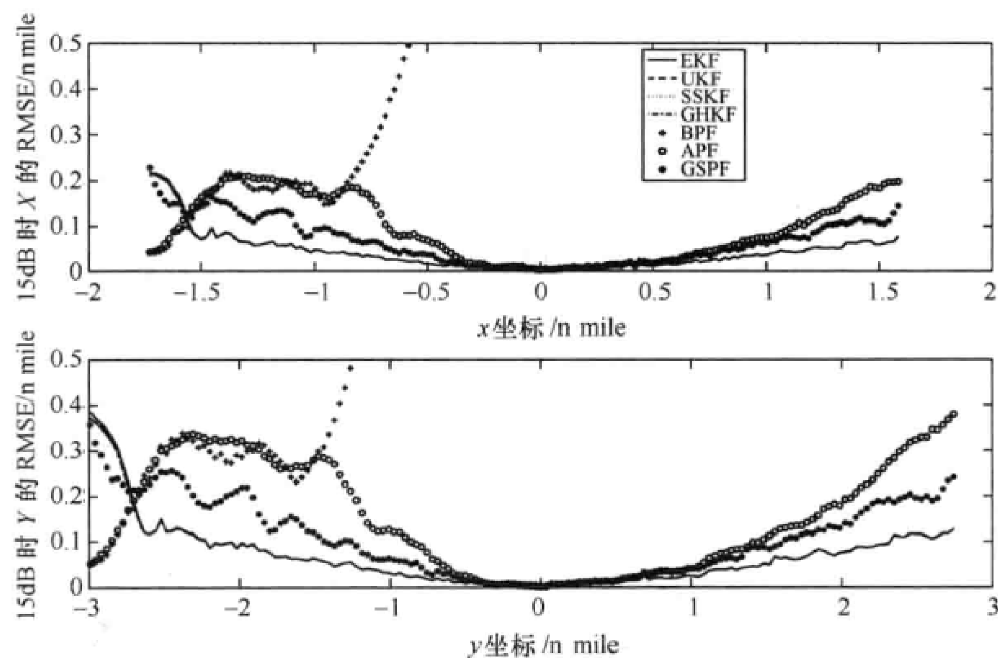


图 17－6　当输入信号信噪比为 15dB 时，DIFAR 实例均方根位置误差比较

图 17－7 给出了 10dB 信号的 RMS 位置误差。在这个信号水平，APF 有 4 个发散航迹（参见图 16－6），导致 RMS 误差增加。但 GSPF 和所有高斯滤波器仍然很稳定，这样它们的 RMS 误差仍然是可以接受的。这里不会给出信噪比低于 10dB 时的 RMS 位置误差图，因为这时所有滤波器的蒙特卡洛跟踪估计都出现一些发散航迹，导致 RMS 误差计算无效。

根据上述性能分析以及第 14 章结尾所指出的，只有当信噪比大于 5dB 时，才能够实现穿过 DIFAR 浮标场船舶的跟踪；否则这些观测值没有利用价值。可能有读者会认为还需要更多的粒子来提高粒子滤波器的性能，但将粒子的数目由 1000 增加到 3000 时，性能几乎没有改善。综上所述，BPF 对初始值非常敏感，使用时需要特殊注意。

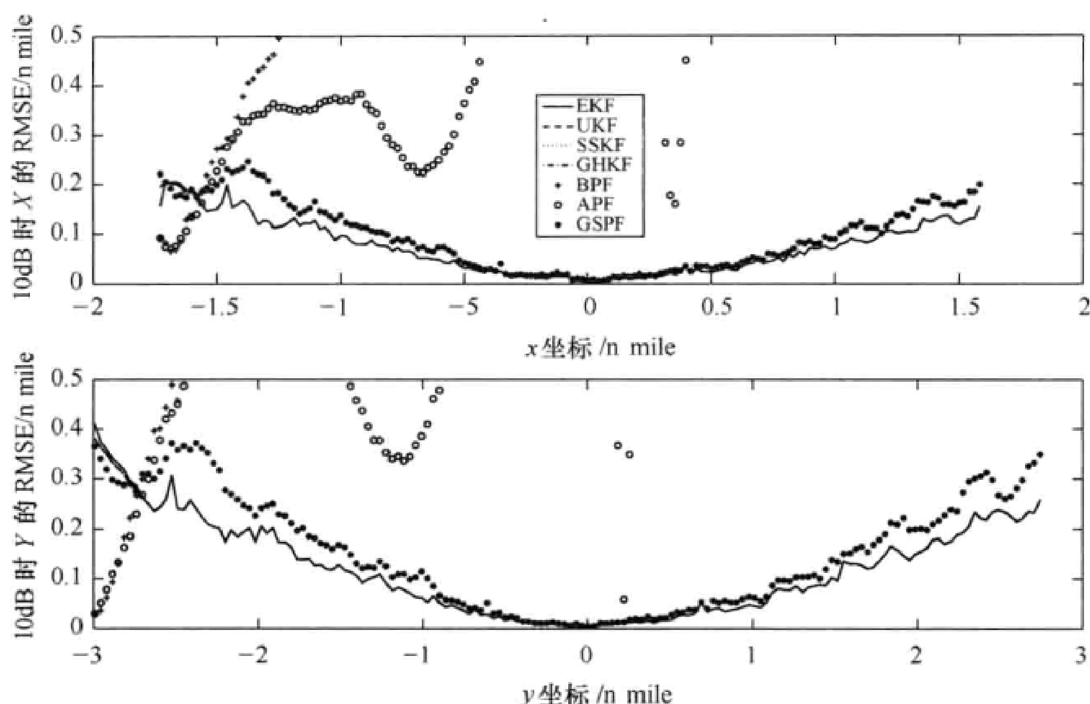


图 17 − 7　当输入信号信噪比为 10dB 时,DIFAR 实例均方根位置误差比较

参考文献

[1] Kotecha JH, Djuric PM. Gaussian Particle Filtering. *Proceedings of the 11th Signal Processing Workshop on Statistical Signal Processing*, 2001:429 − 432.

[2] Kotecha JH, Djuric PM. Gaussian particle filtering. *IEEE Trans. Sig. Proc*, 2003, 51(10):2592 − 2601.

[3] Wan EA, van der Merwe R. The unscented Kalman filter for nonlinear estimation. *Adaptive Systems for Signal Processing, Communications and Control Symposium*, AS − SPEC, 2000:153 − 158.

[4] Kotecha JH, Djuric PM. Gaussian sum particle filtering. *IEEE Trans. Sig. Proc*, 2003, 51(10):2602 − 2612.

[5] van der Merwe R. Wan E. Gaussian Mixture Sigma − Point Particle Filters for Sequential Probabilistic Inference in Dynamic State − Space Models. ICASSP, 2003.

[6] Sornette D, Ide K. The Kalman − Levy filter. *Physica D*, 2001, 115:142 − 174.

[7] Balakrishna N. Statistical signal extraction using stable processes. *Stat. Prob. Lett*, 2009, 79(7):851 − 856.

[8] Wang D, Zhang C, Zhao X. Multivariate Laplace filter: a heavy − tailed model for target tracking. 19th International Conference on Pattern Recognition, 2008:1 − 4.

[9] Gordon N, Percival J, Robinson M. The Kalman − Levy filter and heavy − tailed models for tracking maneuvering targets. *Proceedings of the 6th International Conference on Information Fusion*, 2003:1024 − 1031.

[10] Sinha A. Kirubarafan T. Bar − Shalom Y. Application of the Kalman − Levy filter for tracking maneuvering targets. *Signal Data Process. Small Targets*, 2004, 5428:102 − 112.

[11] Chen Y, Singpurwalla ND. A non − Gaussian Kalman filter model for tracking softwarereliability. Stat. Sinica, 1994, 4(2):535 − 548.

第IV部分

附加案例分析

第 18 章
用于三维目标跟踪的球坐标恒速度模型

这个案例致力于解决一个非常重要的估计问题，也是在许多工程应用的核心问题 —— 三维空间物体跟踪。许多三维跟踪算法利用惯性坐标系下的笛卡儿恒速度线性动态模型来估计目标运动，并用非线性观测模型将笛卡儿坐标状态矢量$\boldsymbol{x}$与一组球坐标观测值z联系起来。如果将动态噪声和观测噪声设为高斯分布，执行卡尔曼滤波器得一组线性卡尔曼滤波器状态预测方程和非线性观测预测方程，可以利用本书第 Ⅱ 部分提出的任何一种次最优滤波技术对方程进行求解。

已经提出多种方法，来研究解决机动（加速）目标跟踪问题。对于笛卡儿坐标跟踪，最简单的方法是在目标动态模型中去掉恒速度的约束，即通过在状态矢量中加入加速度分量，以及用恒加速度约束[1-3]取代速度恒定约束。当目标只有很短的时间是机动的，大部分的航迹上都是非机动情况下，这种算法会出现问题。为处理这类情况，一些学者提出，在检测到机动性时使用滤波器，从恒速度模型切换到恒加速度模型，但机动性的检测并不容易。文献[2 - 4]给出另一种方法，该方法使用很多模型，每个模型用来跟踪不同的潜在的机动目标。还有另一种模型由交互多模型（Interactive Multiple Model，IMM）构成，其中使用一族滤波器模型，模型间只有动态噪声有所不同，其输出构成了混合分布。对 IMM 算法感兴趣的读者可参考上述文献及文献中包括的参考文献。

尽管上述机动目标跟踪方法对于一些场景跟踪效果很好，一些学者也提出了其它的方法，将目标动态模型转换至球坐标系，以保持惯性笛卡儿坐标系零加速度的约束。在文献[5 - 8]提出的方法中，一组正交笛卡儿坐标轴相对于惯性坐标系进行旋转，这样始终有一个轴与连线轴（距离轴）对齐。状态矢量在旋转的球坐标系中，由于旋转坐标系产生科里奥利力（Coriolis forces），必须将伪加速度引入到动态运动方程中。伪加速度加上惯性坐标系笛卡儿恒速度的假设，得

一组非线性完全耦合的球坐标动态方程。当应用于固定的惯性坐标系[5]下跟踪机动目标时，这些滤波器的性能很低。

本章将提出一种更为简单的目标动态模型，适用于沿着机动和非机动航迹跟踪目标。该模型用惯性坐标系球坐标恒速度约束取代惯性坐标系笛卡儿坐标恒速度约束，其中限制正交球速度分量为常数。虽然机动目标经常会违反恒速度约束，但试验表明，在机动过程中，球坐标约束下的 RMS 跟踪误差比较小。文献[5]在以前就给出了类似的结果。

本章后面部分将进行笛卡儿坐标恒速度假设与球坐标恒速度假设跟踪算法性能的比较。因此，从第 18.1 节开始研究在笛卡儿坐标系下目标跟踪，这里假定笛卡儿速度恒定，然后给出动态方程和观测方程，并阐述了笛卡儿坐标系下的扩展卡尔曼滤波器和 Sigma 点卡尔曼滤波器。

第 18.2 节提出一种新的非线性恒定球速度动态方程和线性观测方程，这是根据第一原则导出的。因为动态方程是指一个惯性不旋转坐标系，方程中不会出现伪加速度。基于这个模型，给出了球坐标扩展卡尔曼滤波器和 Sigma 点卡尔曼滤波器。对于远距离情况，可将一般的非线性球坐标动态方程化简为一组线性恒定角速率动态方程，得更为简单的球坐标线性卡尔曼滤波器。

第 18.3 节讨论了滤波器实现问题，第 18.3.1 节起始部分描述了动态噪声项 q(后面将给出其定义）的分量的含义，以及如何根据预期目标加速度限制常识来设置它们的值。所有跟踪滤波器都需要两组必不可少的数据集合作为输入：观测值集合和初始值集合。第 18.3.2 节说明如何生成一个理论上的雷达的观测值，供第 18.4 节中所有的性能评估使用。第 18.3.3 节讨论了根据最开始很少的含噪观测值生成初值的方法。

第 18.4 节比较了笛卡儿坐标系滤波器和球坐标系滤波器的性能。通过对目标机动水平逐渐增加的 3 个仿真航迹的性能分析，可以看到，当目标为机动的且目标距离接近于零时，球坐标跟踪估计滤波器有更好的性能。

第 18.5 节将给出这个案例的一些观测值，并对上述工作未来可能的发展方向进行了展望。

本章结尾给出两个附录，它们对于理解三维跟踪的发展非常重要。因为仿真实例贯穿全章，附录 18.A 讨论了在三维空间中生成机动目标航迹的方法。通常是通过使用高度保真的6 自由度(6DOF）仿真机动对象，利用一个完整的物体作用力模型，如推力、摩擦力、重力等，来实现目标跟踪。因为本书只对不同跟踪算法的之间的比较感兴趣，在附录 18.A 中给出一个低精度的物体运动仿真模型，该模型只能简单地对可能的机动性进行建模，通过将恒定的水平和垂直转弯速率结合起来，基于期望的转弯／俯冲加速度得逼真的航迹。利用这个恒定转弯速率解析模型，生成 3 个航迹实例，实例中机动目标的数目不断增多。第 18.4 节利用这些航迹在目标机动数量增加时，评估跟踪器的性能。

当采用跟踪算法在三维空间跟踪一个物体时,可以选择坐标系统,通常是笛卡儿坐标系或者球坐标系。由于跟踪应用中使用的多数传感器,例如雷达或声纳,所得的测量值都是球坐标系下的,因此6维(3个位置和3个速度分量)坐标变换子程序是必不可少的工具。多数坐标变换子程序只处理状态矢量变换,然而许多跟踪算法需要协方差矩阵坐标变换或需要计算雅可比矩阵。附录18.B给出完整的坐标变换:笛卡儿坐标系到球坐标系变换、球坐标系到笛卡儿坐标系变换以及协方差矩阵的计算。作为协方差变换的附带结果,也会产生雅可比矩阵。因为在分析中使用向东 - 向北 - 向上(x,y,z 轴正方向)笛卡儿坐标系统,也将介绍在这个坐标系统的变换。最后提出了一种通用方法,让读者推导出任意三维坐标系统下的这种变换。

18.1 笛卡儿坐标系下目标跟踪

本节讨论使用定义在笛卡儿坐标系下的状态矢量的进行目标跟踪的方法。本节分为若干小节,讨论用动态模型来定义预期对象运动(18.1.1节)、观测矢量及其如何与状态矢量相关(18.1.2节)、还有多种可以用于解决这个问题的笛卡儿高斯滤波器跟踪算法(18.1.3节)。滤波器初始化和生成蒙特卡洛观测值集合的讨论将会放到第18.3节。

18.1.1 目标动态运动模型

在笛卡儿坐标系下,用一组静止雷达来跟踪运动物体。假设物体以恒定速度运动,可得如下形式的线性动态模型,即

$$\boldsymbol{x}_n = \boldsymbol{F}_{n-1}\boldsymbol{x}_{n-1} + \boldsymbol{v}_{n-1} \tag{18-1}$$

t_n 时刻的笛卡儿坐标系状态矢量定义为

$$\boldsymbol{x}_n \triangleq \boldsymbol{x}(t_n) = [r_{x,n}v_x, r_{y,n}v_y, r_{z,n}, v_z]^{\mathrm{T}} \tag{18-2}$$

转移矩阵 $\boldsymbol{F}_n$ 定义为

$$\boldsymbol{F}_n = \begin{bmatrix} \boldsymbol{F}_{1,n} & \boldsymbol{0}_2 & \boldsymbol{0}_2 \\ \boldsymbol{0}_2 & \boldsymbol{F}_{1,n} & \boldsymbol{0}_2 \\ \boldsymbol{0}_2 & \boldsymbol{0}_2 & \boldsymbol{F}_{1,n} \end{bmatrix} \tag{18-3}$$

$$\boldsymbol{F}_{1,n} = \begin{bmatrix} 1 & T_n \\ 0 & 1 \end{bmatrix} \tag{18-4}$$

$\boldsymbol{0}_2$ 为一个 2×2 的零矩阵,T_n 定义为

$$T_n \triangleq t_n - t_{n-1} \tag{18-5}$$

考虑到传感器跟踪的时间异步性质,可将 T_n 设为时间变量。然而,在后续分析中,只用到单个传感器的情况,所以将 T 设为一个不随时间变化的量。

设动态加速度噪声 $\boldsymbol{v}_n$ 为独立零均值高斯加速度噪声，定义为

$$\boldsymbol{v}_n \backsim \mathcal{N}(\boldsymbol{0},\boldsymbol{Q}_n) \tag{18-6}$$

这里符号 ~ 表示 $\boldsymbol{v}_n$ 来自高斯分布 $\mathcal{N}(\boldsymbol{0},\boldsymbol{Q}_n)$，其中

$$\boldsymbol{Q}_n = \begin{bmatrix} q_{x,n}\boldsymbol{Q}_1 & \boldsymbol{0}_2 & \boldsymbol{0}_2 \\ \boldsymbol{0}_2 & q_{y,n}\boldsymbol{Q}_1 & \boldsymbol{0}_2 \\ \boldsymbol{0}_2 & \boldsymbol{0}_2 & q_{z,n}\boldsymbol{Q}_1 \end{bmatrix} \tag{18-7}$$

对于连续运动对象，$\boldsymbol{Q}_1$ 由文献[3]给出，即

$$\boldsymbol{Q}_1 = \begin{bmatrix} \frac{T^3}{3} & \frac{T^2}{2} \\ \frac{T^2}{2} & T \end{bmatrix} \tag{18-8}$$

$q_{i,n}\ \forall i$ 是一个重要的滤波器设计参数。当研究各种滤波器设计选择时，基于期望的物体运动的特点（例如，对象是机动还是直线运动），给出如何选择 $q_{i,n}$。

18.1.2 传感器数据模型

正如 18.3.2 节指出的，多数雷达传感器的输出观测（测量）数据是二维极坐标或三维球坐标，而与动态模型所选的坐标系没有关系。因此，将完整的传感器数据矢量定义为

$$z_n = [R_n, \dot{R}_n, \theta_n, \phi_n]^{\mathrm{T}} \tag{18-9}$$

其中集合 $\{R_n, \dot{R}_n, \theta_n, \phi_n\}$ 分别为待跟踪目标相对于固定传感器位置的距离、多普勒（距离变化率）、方位和高程。对于那些没有完整观测分量的传感器，只能放弃那些对于特定雷达无法获取的分量。例如，在后面用于性能分析的仿真中，所选的传感器模型中只包含距离、方位和高程。

通过后面式(18-160)和式(18-161)定义的非线性球坐标到笛卡儿坐标变换，得观测矢量与笛卡儿状态矢量之间关系，即

$$z_n = \begin{bmatrix} R_n \\ \dot{R}_n \\ \theta_n \\ \phi_n \end{bmatrix} = \boldsymbol{h}_n(\boldsymbol{x}_n) = \begin{bmatrix} h_{1,n}(\boldsymbol{x}_n) \\ h_{2,n}(\boldsymbol{x}_n) \\ h_{3,n}(\boldsymbol{x}_n) \\ h_{4,n}(\boldsymbol{x}_n) \end{bmatrix} = \begin{bmatrix} \sqrt{r_x^2 + r_y^2 + r_z^2} \\ \dfrac{r_x v_x + r_y v_y + r_z v_z}{R} \\ \tan^{-1}\left(\dfrac{r_x}{r_y}\right) \\ \tan^{-1}\left(\dfrac{r_z}{\sqrt{r_x^2 + r_y^2}}\right) \end{bmatrix} \tag{18-10}$$

通常传感器噪声包含在观测模型中，因此观测模型变为

$$\boldsymbol{z}_n = \boldsymbol{h}_n(\boldsymbol{x}_n) + \boldsymbol{w}_n \tag{18-11}$$

观测噪声 $\boldsymbol{w}_n$ 与零均值高斯噪声无关，定义为

$$\boldsymbol{w}_n \backsim \mathcal{N}(\boldsymbol{0},\boldsymbol{R}_n) \tag{18-12}$$

$\boldsymbol{R}_n$ 通常认为与时间无关,可以用传感器测量值方差的对角协方差矩阵定义为

$$\boldsymbol{R} = \begin{bmatrix} \sigma_R^2 & 0 & 0 & 0 \\ 0 & \sigma_R^2 & 0 & 0 \\ 0 & 0 & \sigma_\theta^2 & 0 \\ 0 & 0 & 0 & \sigma_\phi^2 \end{bmatrix} \tag{18-13}$$

对于真实传感器数据,在每一个时间传感器更新时,通常作为传感器数据流的一部分输出传感器测量标准偏差。如果传感探测器的输出信噪比不断变化,所以 $\boldsymbol{R}$ 的分量应该是时间相关的。对于以下性能分析用到的仿真,设 $\boldsymbol{R}$ 为不随时间变化的。

18.1.3 高斯跟踪算法计算笛卡儿状态矢量

18.1.3.1 笛卡儿线性状态预测方程

因为笛卡儿动态方程是线性的,可以使用表6-1中的线性状态方程和状态协方差预测方程。

$$\hat{\boldsymbol{x}}_{n|n-1} = \boldsymbol{F}_{n-1}\hat{\boldsymbol{x}}_{n-1|n-1} \tag{18-14}$$

$$\boldsymbol{P}^{xx}_{n|n-1} = \boldsymbol{F}_{n-1}\boldsymbol{P}^{xx}_{n-1|n-1}\boldsymbol{F}^{\mathrm{T}}_{n-1} + \boldsymbol{Q}_{n-1} \tag{18-15}$$

18.1.3.2 笛卡儿非线性观测预测

由式(18-10)可见,观测模型关于状态矢量是非线性的,所以必须用第Ⅱ部分的一种次优滤波器进行观测预测。在这个实例中,仅使用 EKF 和除 MCKF 之外所有的 Sigma 点卡尔曼滤波器。

18.1.3.2.1 笛卡儿线性化 EKF 观测预测方程。根据表7-2,EKF 观测预测方程为

$$\hat{\boldsymbol{z}}_{n|n-1} = \boldsymbol{h}_n(\hat{\boldsymbol{x}}_{n|n-1}) \tag{18-16}$$

$$\boldsymbol{P}^{zz}_{n|n-1} = \hat{\boldsymbol{H}}_{n|n-1}\boldsymbol{P}^{xx}_{n|n-1}\hat{\boldsymbol{H}}^{\mathrm{T}}_{n|n-1} + \boldsymbol{R}_n \tag{18-17}$$

$$\boldsymbol{P}^{xz}_{n|n-1} = \boldsymbol{P}^{xx}_{n|n-1}\hat{\boldsymbol{H}}^{\mathrm{T}}_{n|n-1} \tag{18-18}$$

其中式(18-164)~式(18-190)给出了雅可比矩阵的元素 $\hat{\boldsymbol{H}}_{n|n-1}$。

18.1.3.2.2 笛卡儿非线性 Sigma 点观测预测方程。根据图13-3,给出 Sigma 点卡尔曼滤波器预测方程为

$$\boldsymbol{D}_{n|n-1} = [\boldsymbol{P}^{xx}_{n|n-1}]^{1/2} \tag{18-19}$$

$$\boldsymbol{\chi}^{(j)}_{n|n-1} = \hat{\boldsymbol{x}}_{n|n-1} + \boldsymbol{D}_{n|n-1}\boldsymbol{c}^{(j)} \tag{18-20}$$

$$\boldsymbol{\psi}_{n|n-1}^{(j)} = \boldsymbol{h}_n(\boldsymbol{\chi}_{n|n-1}^{(j)}) \tag{18-21}$$

$$\hat{\boldsymbol{z}}_{n|n-1} = \sum_{j=1}^{N_s} w_j \boldsymbol{\psi}_{n|n-1}^{(j)} \tag{18-22}$$

$$\boldsymbol{P}_{n|n-1}^{zz} = \sum_{j=1}^{N_s} w_j (\boldsymbol{\psi}_{n|n-1}^{(j)} - \hat{\boldsymbol{z}}_{n|n-1})(\boldsymbol{\psi}_{n|n-1}^{(j)} - \hat{\boldsymbol{z}}_{n|n-1})^{\mathrm{T}} + \boldsymbol{R}_{n-1} \tag{18-23}$$

$$\boldsymbol{P}_{n|n-1}^{xz} = \sum_{j=1}^{N_s} w_j (\boldsymbol{\chi}_{n|n-1}^{(j)} - \hat{\boldsymbol{x}}_{n|n-1})(\boldsymbol{\psi}_{n|n-1}^{(j)} - \hat{\boldsymbol{z}}_{n|n-1})^{\mathrm{T}} \tag{18-24}$$

这里集合$\{w_j, \boldsymbol{c}^{(j)}\}$可以用列表 13 - 1 和表 13 - 2 离线计算(参见第 13 章)。

18.1.3.3 卡尔曼滤波器更新方程

跟踪估计的最后步骤是使用卡尔曼滤波器更新方程(5 - 29) - 方程(5 - 31),即

$$\boldsymbol{K}_n = \boldsymbol{P}_{n|n-1}^{xz} [\boldsymbol{P}_{n|n-1}^{zz}]^{-1} \tag{18-25}$$

$$\hat{\boldsymbol{x}}_{n|n} = \hat{\boldsymbol{x}}_{n|n-1} + \boldsymbol{K}_n (z_n - \hat{z}_{n|n-1}) \tag{18-26}$$

$$\boldsymbol{P}_{n|n-1}^{xx} = \boldsymbol{P}_{n|n-1}^{xx} - \boldsymbol{K}_n \boldsymbol{P}_{n|n-1}^{zz} \boldsymbol{K}_n^{\mathrm{T}} \tag{18-27}$$

18.2 球坐标系下目标跟踪

本节将讨论利用球坐标系下状态矢量进行跟踪目标,状态矢量定义为

$$\rho = [R, v_r, \theta, v_h, \phi, v_v]^{\mathrm{T}} \tag{18-28}$$

其中球速度分量$\{v_r, v_h, v_v\}$将在后面介绍。与笛卡儿坐标系下情况类似,假定笛卡儿速度分量为常数,导出式(18 - 1)的笛卡儿坐标系动态运动模型,这里研究目标是建立具有恒定球速度分量的目标的三维动态运动模型。

18.2.1 球坐标系下状态矢量的位置和速度分量

参考图 18 - 25,目标的位置可写为

$$\boldsymbol{r}_p = x\boldsymbol{e}_x + y\boldsymbol{e}_y + z\boldsymbol{e}_z \tag{18-29}$$

$$\boldsymbol{r}_p = a_r\boldsymbol{e}_r + a_\theta\boldsymbol{e}_\theta + a_\phi\boldsymbol{e}_\phi \tag{18-30}$$

其中,$\{\boldsymbol{e}_x, \boldsymbol{e}_y, \boldsymbol{e}_z\}$和$\{\boldsymbol{e}_R, \boldsymbol{e}_\theta, \boldsymbol{e}_\phi\}$为分别在笛卡儿坐标系和球坐标系下的单位矢量集合。

位置分量从球坐标系到笛卡儿坐标系的变换矩阵由式(18 - 193)定义,所以可以将式(18 - 29)改写为

$$\boldsymbol{r}_p = \boldsymbol{e}_x\boldsymbol{R}\sin\theta\cos\phi + \boldsymbol{e}_y\boldsymbol{R}\cos\theta\cos\phi + \boldsymbol{e}_z\boldsymbol{R}\sin\phi \tag{18-31}$$

根据定义,球坐标系单位矢量为

$$\boldsymbol{e}_r = \frac{1}{\left|\dfrac{\partial \boldsymbol{r}_p}{R}\right|} \frac{\partial \boldsymbol{r}_p}{R} \tag{18-32}$$

$$e_\theta = \frac{1}{\left|\frac{\partial \boldsymbol{r}_p}{\theta}\right|}\frac{\partial \boldsymbol{r}_p}{\theta} \tag{18-33}$$

$$e_\phi = \frac{1}{\left|\frac{\partial \boldsymbol{r}_p}{\phi}\right|}\frac{\partial \boldsymbol{r}_p}{\phi} \tag{18-34}$$

将式(18－31)代入式(18－32)，根据集合$\{\boldsymbol{e}_x, \boldsymbol{e}_y, \boldsymbol{e}_z\}$得$\boldsymbol{e}_r$的表达式为

$$\boldsymbol{e}_r = \boldsymbol{e}_x \sin\theta\cos\phi + \boldsymbol{e}_y \cos\theta\cos\phi + \boldsymbol{e}_z \sin\phi \tag{18-35}$$

类似地，得

$$\boldsymbol{e}_\theta = \boldsymbol{e}_x \cos\theta - \boldsymbol{e}_y \sin\theta \tag{18-36}$$

$$\boldsymbol{e}_\phi = -\boldsymbol{e}_x \sin\theta\sin\phi - \boldsymbol{e}_y \cos\theta\sin\phi + \boldsymbol{e}_z \cos\phi \tag{18-37}$$

以下给出了球面单位矢量的时间导数，即

$$\begin{aligned}\dot{\boldsymbol{e}}_r &= \frac{\partial \boldsymbol{e}_r}{\partial R}\dot{R} + \frac{\partial \boldsymbol{e}_r}{\partial \theta}\dot{\theta} + \frac{\partial \boldsymbol{e}_r}{\partial \phi}\dot{\phi} \\ &= \boldsymbol{e}_\theta \dot{\theta}\cos\phi + \boldsymbol{e}_\phi \dot{\phi}\end{aligned} \tag{18-38}$$

$$\begin{aligned}\dot{\boldsymbol{e}}_\theta &= \frac{\partial \boldsymbol{e}_\theta}{\partial R}\dot{R} + \frac{\partial \boldsymbol{e}_\theta}{\partial \theta}\dot{\theta} + \frac{\partial \boldsymbol{e}_\theta}{\partial \phi}\dot{\phi} \\ &= -\boldsymbol{e}_r \dot{\theta}\cos\phi + \boldsymbol{e}_\phi \dot{\theta}\sin\phi\end{aligned} \tag{18-39}$$

$$\begin{aligned}\dot{\boldsymbol{e}}_\phi &= \frac{\partial \boldsymbol{e}_\phi}{\partial R}\dot{R} + \frac{\partial \boldsymbol{e}_\phi}{\partial \theta}\dot{\theta} + \frac{\partial \boldsymbol{e}_\phi}{\partial \phi}\dot{\phi} \\ &= -\boldsymbol{e}_\theta \dot{\theta}\sin\phi - \boldsymbol{e}_r \dot{\phi}\end{aligned} \tag{18-40}$$

利用式(18－35)，可以将式(18－31)写成

$$\begin{aligned}\boldsymbol{r}_p &= \boldsymbol{e}_x R\sin\theta\cos\phi + \boldsymbol{e}_y R\cos\theta\cos\phi + \boldsymbol{e}_z R\sin\phi \\ &= R[\boldsymbol{e}_x \sin\theta\cos\phi + \boldsymbol{e}_y \cos\theta\cos\phi + \boldsymbol{e}_z \sin\phi] \\ &= R\boldsymbol{e}_r\end{aligned} \tag{18-41}$$

由于速度矢量是位置矢量关于时间的导数，得

$$\begin{aligned}\boldsymbol{v} &= \frac{\mathrm{d}}{\mathrm{d}t}(R\boldsymbol{e}_r) = \dot{R}\boldsymbol{e}_r + R\dot{\boldsymbol{e}}_r \\ &= \boldsymbol{e}_r \dot{R} + \boldsymbol{e}_\theta R\dot{\theta}\cos\phi + \boldsymbol{e}_\phi R\dot{\phi}\end{aligned} \tag{18-42}$$

确定球坐标系中的分量$\boldsymbol{v}$为

$$\boldsymbol{v} = \begin{bmatrix} v_r \\ v_h \\ v_v \end{bmatrix} = \begin{bmatrix} \dot{R} \\ R\dot{\theta}\cos\phi \\ R\dot{\phi} \end{bmatrix} = \begin{bmatrix} \dot{R} \\ R_h\dot{\theta} \\ R\dot{\phi} \end{bmatrix} \tag{18-43}$$

这里定义水平距离为$R_h \triangleq R\cos\phi$。利用式(18－43)作为速度分量，式(18－28)给出了球坐标系状态矢量的所有分量。

18.2.2 球坐标系状态矢量动态方程

对于恒速度目标运动动态模型,假设 $\dot{v}_r = \dot{v}_h = \dot{v}_v = 0$。取式(18 - 43)各个分量关于时间的导数,得

$$\dot{v}_r = \ddot{R} = 0 \tag{18 - 44}$$

$$\dot{v}_h = \dot{R}\dot{\theta}\cos\phi + R\ddot{\theta}\cos\phi - R\dot{\theta}\dot{\phi}\sin\phi = 0 \tag{18 - 45}$$

$$\dot{v}_v = \dot{R}\dot{\phi} + R\ddot{\phi} = 0 \tag{18 - 46}$$

或者,重新整理,得

$$\ddot{\theta} = \dot{\theta}\left(\dot{\phi}\tan\phi - \frac{\dot{R}}{R}\right) \tag{18 - 47}$$

$$\ddot{\phi} = -\frac{\dot{R}}{R}\dot{\phi} \tag{18 - 48}$$

下面将给出动态运动方程的分量。距离可以写成

$$R_n = R_{n-1} + T\dot{R}_{n-1} + \frac{T^2}{2}\ddot{R}_{n-1} \tag{18 - 49}$$

将式(18 - 44)和式(18 - 43)中 v_r 的定义代入上式,变为

$$R_n = R_{n-1} + Tv_{r,n-1} \tag{18 - 50}$$

对于径向速度,有

$$v_{r,n} = v_{r,n-1} + T\dot{v}_r = v_{r,n-1} \tag{18 - 51}$$

用类似于式(18 - 49)的形式写出方位方程,用式(18 - 43)速度分量的定义取代式(18 - 47)中的 $\ddot{\theta}_{n-1}$,得

$$\theta_n = \theta_{n-1} + T\frac{v_{h,n-1}}{R_{h,n-1}}\left[1 + \frac{T}{2R_{n-1}}(v_{v,n-1}\text{tna}\phi_{n-1} - v_{r,n-1})\right] \tag{18 - 52}$$

与式(18 - 51)形式类似,有

$$v_{h,n} = v_{h,n-1} \tag{18 - 53}$$

此外,高程分量为

$$\phi_n = \phi_{n-1} + T\frac{v_{v,n-1}}{R_{n-1}}\left[1 - \frac{T}{2}\frac{v_{r,n-1}}{R_{n-1}}\right] \tag{18 - 54}$$

$$v_{v,n} = v_{v,n-1} \tag{18 - 55}$$

球坐标状态矢量的非线性动态运动方程可以写成

$$\boldsymbol{\rho}_n = \boldsymbol{f}(\boldsymbol{\rho}_{n-1}) + \boldsymbol{n}_{n-1} \tag{18 - 56}$$

其中零均值高斯动态加速度噪声分布为

$$\boldsymbol{n}_n \backsim \mathcal{N}(\boldsymbol{0}, \boldsymbol{Q}_\rho) \tag{18 - 57}$$

$\boldsymbol{f}(\boldsymbol{\rho}_{n-1})$ 的元素由非线性方程(18 - 50) ~ 式(18 - 55)以及下式给出,即

$$\boldsymbol{Q}_\rho = \begin{bmatrix} q_r\boldsymbol{Q}_1 & \boldsymbol{0}_2 & \boldsymbol{0}_2 \\ \boldsymbol{0}_2 & q_h\boldsymbol{Q}_1 & \boldsymbol{0}_2 \\ \boldsymbol{0}_2 & \boldsymbol{0}_2 & q_v\boldsymbol{Q}_1 \end{bmatrix} \tag{18-58}$$

$$\boldsymbol{Q}_1 = \begin{bmatrix} \frac{T^3}{3} & \frac{T^2}{2} \\ \frac{T^2}{2} & T \end{bmatrix} \tag{18-59}$$

18.2.2.1 非线性球坐标动态方程的远程限制

通过使用式(18－43)，分别利用$\dot{\theta}_{n-1}$和$\dot{\phi}_{n-1}$替换式(18－52)和式(18－54)中的$v_{h,n-1}/R_{h,n-1}$和$v_{v,n-1}/R_{n-1}$。限定$R_{n-1}\to\infty$，式(18－50)～式(18－55)变为线性恒速率动态方程，即

$$R_n = R_{n-1} + T\dot{R}_{n-1} \tag{18-60}$$

$$\dot{R}_n = \dot{R}_{n-1} \tag{18-61}$$

$$\theta_n = \theta_{n-1} + T\dot{\theta}_{n-1} \tag{18-62}$$

$$\dot{\theta}_n = \dot{\theta}_{n-1} \tag{18-63}$$

$$\phi_n = \phi_{n-1} + T\dot{\phi}_{n-1} \tag{18-64}$$

$$\dot{\phi}_n = \dot{\phi}_{n-1} \tag{18-65}$$

这种情况下，t_n时刻的球坐标状态矢量可以表示为

$$\boldsymbol{r}_n = [R_n, \dot{R}_n, \theta_n, \dot{\theta}_n, \phi_n, \dot{\phi}_n]^{\mathrm{T}} \tag{18-66}$$

动态方程变为线性方程，即

$$\boldsymbol{r}_n = \boldsymbol{F}\boldsymbol{r}_{n-1} + \boldsymbol{n}_{n-1} \tag{18-67}$$

$$\boldsymbol{F} = \begin{bmatrix} \boldsymbol{F}_1 & \boldsymbol{0}_2 & \boldsymbol{0}_2 \\ \boldsymbol{0}_2 & \boldsymbol{F}_1 & \boldsymbol{0}_2 \\ \boldsymbol{0}_2 & \boldsymbol{0}_2 & \boldsymbol{F}_1 \end{bmatrix} \tag{18-68}$$

$$\boldsymbol{F}_1 = \begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix} \tag{18-69}$$

$$\boldsymbol{Q}_r = \begin{bmatrix} q_r\boldsymbol{Q}_1 & \boldsymbol{0}_2 & \boldsymbol{0}_2 \\ \boldsymbol{0}_2 & q_\theta\boldsymbol{Q}_1 & \boldsymbol{0}_2 \\ \boldsymbol{0}_2 & \boldsymbol{0}_2 & q_\phi\boldsymbol{Q}_1 \end{bmatrix} \tag{18-70}$$

18.2.3 球坐标系状态矢量观测方程

式(18－9)给出了传感器数据矢量。基于这个数据矢量，可写出与式(18－56)

或式(18 - 67)球坐标动力学方程相关的线性观测方程，即

$$z_n = \boldsymbol{H}\boldsymbol{\rho}_n + \boldsymbol{w}_{n-1} \tag{18-71}$$

$$= \boldsymbol{H}\boldsymbol{r}_n + \boldsymbol{w}_{n-1} \tag{18-72}$$

$$\boldsymbol{H} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \tag{18-73}$$

18.2.4 高斯跟踪算法计算球坐标状态矢量

18.2.4.1 当 R 很大时的跟踪估计

首先考虑第一种 $R \gg 1$ 的情况，可以看到，动态方程和观测方程即式(18 - 67)和式(18 - 72)都是线性的，因此第6章提出的线性卡尔曼滤波器是一个最优滤波器。第6章给出这种情况的 LKF 状态预测方程为

$$\hat{\boldsymbol{r}}_{n|n-1} = \boldsymbol{F}\hat{\boldsymbol{r}}_{n-1|n-1} \tag{18-74}$$

$$\boldsymbol{P}^{rr}_{n|n-1} = \boldsymbol{F}\boldsymbol{P}^{rr}_{n-1|n-1}\boldsymbol{F}^{\mathrm{T}} + \boldsymbol{Q}_r \tag{18-75}$$

$\boldsymbol{F}$ 和 $\boldsymbol{Q}$ 分别由式(18 - 68)和式(18 - 70)给出。

观测预测方程可以写成

$$\hat{z}_{n|n-1} = \boldsymbol{H}\hat{\boldsymbol{r}}_{n|n-1} \tag{18-76}$$

$$\boldsymbol{P}^{zz}_{n|n-1} = \boldsymbol{H}\boldsymbol{P}^{rr}_{n|n-1}\boldsymbol{H}^{\mathrm{T}} + \boldsymbol{R} \tag{18-77}$$

$$\boldsymbol{P}^{xz}_{n|n-1} = \boldsymbol{P}^{rr}_{n|n-1}\boldsymbol{H}^{\mathrm{T}} \tag{18-78}$$

$\boldsymbol{H}$ 和 $\boldsymbol{R}$ 分别由式(18 - 73)和式(18 - 13)给出。

卡尔曼滤波器更新方程可写成

$$\boldsymbol{K}_n = \boldsymbol{P}^{zz}_{n|n-1}\left[\boldsymbol{P}^{zz}_{n|n-1}\right]^{-1} \tag{18-79}$$

$$\hat{\boldsymbol{r}}_{n|n} = \hat{\boldsymbol{r}}_{n|n-1} + \boldsymbol{K}_n(z_n - \hat{z}_{n|n-1}) \tag{18-80}$$

$$\boldsymbol{P}^{rr}_{n|n} = \boldsymbol{P}^{rr}_{n|n-1} - \boldsymbol{K}_n\boldsymbol{P}^{zz}_{n|n-1}\boldsymbol{K}_n^{\mathrm{T}} \tag{18-81}$$

18.2.4.2 非线性球坐标动态模型的跟踪估计

球坐标系状态矢量的实例与笛卡儿坐标状态矢量的实例正好相反。对于以恒定笛卡儿速度运动的笛卡儿状态矢量，其状态预测方程是线性的，但观测方程是非线性的，需要用次优非线性跟踪滤波器进行观测值预测。现在，状态矢量预测方程是非线性的，而观测预测方程是线性的，需要一个次优非线性跟踪滤波器进行状态预测。下面，首先研究在状态预测步骤使用线性卡尔曼滤波器所需要进行的改进，然后考虑 Sigma 点方法。

18.2.4.2.1 在球坐标系中应用 EKF。为使用 EKF 在球面坐标系中进行状态预

测,必须首先找到非线性状态转移方程的雅可比矩阵,即

$$\boldsymbol{\rho}_n = f(\rho_{n-1}) \tag{18-82}$$

其中式(18-50)~式(18-55)中给出转移方程的元素。这样,EKF状态预测方程为

$$\hat{\boldsymbol{\rho}}_{n|n-1} = f(\hat{\boldsymbol{\rho}}_{n-1|n-1}) \tag{18-83}$$

$$\boldsymbol{P}^{\rho\rho}_{n|n-1} = \hat{\boldsymbol{F}}_{\rho,n-1}\boldsymbol{P}^{\rho\rho}_{n-1|n-1}\hat{\boldsymbol{F}}^{\mathrm{T}}_{\rho,n-1} + \boldsymbol{Q}_\rho \tag{18-84}$$

这里 $\hat{\boldsymbol{F}}_{\rho,n-1}$ 是转移方程(18-82)在 t_{n-1} 时刻的雅可比矩阵。

采用附录18.B的一般方法,式(18-82)的雅可比矩阵可由下式给出,即

$$\hat{\boldsymbol{F}}_\rho = \begin{bmatrix} 1 & T & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ F_{31} & F_{32} & 1 & F_{34} & F_{35} & F_{36} \\ 0 & 0 & 0 & 1 & 0 & 0 \\ F_{51} & F_{52} & 0 & 0 & 1 & F_{56} \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{18-85}$$

$$F_{31} = -\frac{Tv_h}{RR_h}\left[1 + \frac{T}{R}(v_v\tan\phi - v_r)\right] \tag{18-86}$$

$$F_{32} = -\frac{T^2v_h}{2RR_h} \tag{18-87}$$

$$F_{34} = \frac{T}{R_h}\left[1 + \frac{T}{2R}(v_v\tan\phi - v_r)\right] \tag{18-88}$$

$$F_{35} = \frac{T^2v_h}{2RR_h}v_v\sec^2\phi \tag{18-89}$$

$$F_{36} = \frac{T^2v_h}{2RR_h}\tan\phi \tag{18-90}$$

$$F_{51} = -\frac{Tv_v}{R^2}\left[1 + \frac{Tv_r}{R}\right] \tag{18-91}$$

$$F_{52} = \frac{T^2v_v}{2R^2} \tag{18-92}$$

$$F_{56} = \frac{T}{R}\left[1 + \frac{Tv_r}{R}\right] \tag{18-93}$$

这里为清晰起见,移除了下标 n。

因为观测矢量中没有与 v_h 或者 v_v 相关的元素,观测预测方程和更新方程与18.2.4.1节给出的球坐标LKF的方程相同。

18.2.4.2.2 在球坐标系中应用Sigma点卡尔曼滤波器。流程的第一步是根据前面滤波器的输出 $\{\hat{\boldsymbol{\rho}}_{n-1|n-1}, \boldsymbol{P}^{\rho\rho}_{n-1|n-1}\}$ 生成一组Sigma点,或者由初始值利用下式

$$\Pi_{n-1|n-1}^{(j)} = \hat{\boldsymbol{\rho}}_{n-1|n-1} + \boldsymbol{D}_{n-1|n-1}\boldsymbol{c}^{(j)} \quad j = 0,1,\cdots,N_s \tag{18-94}$$

$$\boldsymbol{D}_{n-1|n-1} = [\boldsymbol{P}_{n-1|n-1}^{\rho\rho}]^{1/2} \tag{18-95}$$

用列表 13－1 和 13－2 中给出的子程序来生成集合$\{w_j, \boldsymbol{c}^{(j)}, j = 0,1,\cdots,N_s\}$。它们将取决于特定的 Sigma 点卡尔曼滤波器的选择和状态矢量的维数。

下面给出 Sigma 点卡尔曼滤波器状态预测方程为

$$\Pi_{n|n-1}^{(j)} = \boldsymbol{f}(\Pi_{n-1|n-1}^{(j)}) \tag{18-96}$$

$$\boldsymbol{\rho}_{n|n-1} = \sum_{j=0}^{N_s} w_j \Pi_{n|n-1}^{(j)} \tag{18-97}$$

$$\boldsymbol{P}_{n|n-1}^{\rho\rho} = \sum_{j=0}^{N_s} w_j (\Pi_{n|n-1}^{(j)} - \hat{\boldsymbol{\rho}}_{n|n-1})(\Pi_{n|n-1}^{(j)} - \hat{\boldsymbol{\rho}}_{n|n-1})^{\mathrm{T}} \tag{18-98}$$

$$+ \boldsymbol{Q}_{\rho} \tag{18-99}$$

与球坐标系 EKF 一样，线性观测预测方程和更新方程与上面给出的球坐标 LKF 的方程相同。

18.3 笛卡儿坐标系和球坐标系下跟踪滤波器的实现

为了实现前面提出的跟踪滤波器，必须要

(1) 指定一组 q 值；

(2) 生成含噪数据的蒙特卡洛集合作为滤波器的观测值；

(3) 生成初始化状态矢量和协方差矩阵。

以下各节逐一解决这几个问题。

18.3.1 设置 q 值

根据式(18－15)的笛卡儿状态协方差矩阵的预测方程，如果将式(18－15)等式左边元素的左上角 2×2 集合与类似的由式(18－8)得的元素集合 $\boldsymbol{Q}$ 进行比较，得

$$\begin{bmatrix} P_{xx} & P_{x\dot{x}} \\ P_{\dot{x}x} & P_{\dot{x}\dot{x}} \end{bmatrix} = \begin{bmatrix} \sigma_x^2 & \bullet \\ \bullet & \sigma_{v_x}^2 \end{bmatrix} \propto \begin{bmatrix} q_x \dfrac{T^3}{3} & q_x \dfrac{T^2}{2} \\ q_x \dfrac{T^2}{2} & q_x T \end{bmatrix} \tag{18-100}$$

因此 $\sigma_{v_x}^2 \propto q_x T$，$q_x$ 的单位是$[l^2/t^3]$。或者，可以写成$[a^2 \times t]$ 或者$[a^2/f]$。

假设已知试图跟踪目标的某些机动特征，可以估算出所需 q_x 值。因为动态模型是零均值高斯加速度噪声的恒速度模型，可以设置为目标最大可能机动加速度的 2 倍。它增加了一倍来表示向左或向右转。所以设

$$q_x = (2a)^2 \tag{18-101}$$

如果用 g(重力加速度) 来表示 a，必须将单位变换为$(\mathrm{n\ mile/s^2})^2$。例如，如

果一个目标最大转弯加速度为 $3g$,则有

$$q_x = (2 \times 3g)^2 \times (32.174\text{ft/s}^2/g)^2 \times (1.65 \times 10^{-4}\text{n mile/ft})^2$$
$$= 35 \times 2.80 \times 10^{-5}(\text{n mile/s}^2)^2 \tag{18-102}$$

由式(18－7)可知,q 的各个分量可以设置为不同值。因此,对于一个笛卡儿滤波器,可以将水平机动值(q_x 和 q_y)设定为一个值,而将垂直机动(q_z)设定为另一个值。对于球坐标滤波器,用集合$\{q_r, q_h, q_v\}$表示恒球速度滤波器,用$\{q_r, q_\theta, q_\phi\}$表示恒球速率滤波器。一般来说,可以设 q_r 和 q_h 与笛卡儿水平 q 值相等,设置 q_v 的值等于笛卡儿垂直 q_z。对于恒球速率滤波器,很容易看出 q 的角度可以设为 $q_\theta = q_r/R_h^2$ 和 $q_\phi = q_r/R^2$。后面将看到,球坐标下 q 值越高,性能越好。

18.3.2 雷达观测数据仿真

为对跟踪估计算法进行性能比较,对每种跟踪算法必须在给定的蒙特卡洛次数上用相同的含噪观测值。对于每次蒙特卡洛运行,必须用一组新的、与前面设置无关的含噪数据生成一组不同的观测值。所以,如果要使用 100 次蒙特卡洛运行来比较多目标跟踪估计算法的性能,则需要 100 组含噪观测值,噪声与时间采样无关,也和蒙特卡洛集合无关。

跟踪目标并生成航迹,利用附录 18. A 给出的低精度仿真。忽略大气往返传播对雷达信号以及雷达探测器检测统计数据的影响,假定探测概率为 100%。生成观测集的特征,雷达必须设置:

(1) 雷达观测矢量由哪些变量构成;

(2) 什么是雷达更新率;

(3) 每个观测变量的观测噪声标准偏差是多少。

对于一个普通雷达,其观测矢量可以由目标的距离(R)和距离变化的速率($\dot{R}$)构成,目标相对于正北向的水平方位(θ)和目标在地面以上的高度(ϕ)。这样,观测矢量可写成

$$z = [R, \dot{R}, \theta, \phi]^{\mathrm{T}} \tag{18-103}$$

注意:在球极坐标系中 $-180° \leqslant \theta \leqslant 180°$ 且 $0° \leqslant \phi \leqslant 90°$。

通常假定观测矢量的每个元素的噪声为高斯分布且无关,因此得观测噪声协方差矩阵为

$$\boldsymbol{R} = \begin{bmatrix} \sigma_R^2 & 0 & 0 & 0 \\ 0 & \sigma_{\dot{R}}^2 & 0 & 0 \\ 0 & 0 & \sigma_\theta^2 & 0 \\ 0 & 0 & 0 & \sigma_\phi^2 \end{bmatrix} \tag{18-104}$$

式中:σ_i 为第 i 个观测矢量元素噪声的标准偏差。

在下一节给出的跟踪估计性能结果中,采用一个只有距离、方位和高程的雷

达模型。用于生成含噪观测值的观测不确定性为

$$\sigma_R = 0.15\text{n mile} \quad (18-105)$$

$$\sigma_\theta = 0.3° \quad (18-106)$$

$$\sigma_\phi = 0.3° \quad (18-107)$$

仿真雷达的更新周期为每4s一个观测值,更新频率为0.25Hz。

利用附录18.A中给出的方法,可生成笛卡儿目标航迹,以10Hz的频率生成仿真航迹数据。因为雷达模型的更新频率低得多,航迹数据必须首先降低采样至雷达更新频率,以得添加观测噪声之前的观测值的真实值。当然也可以增加一个开始时间变量,但对于这里的目标来说是没有必要的,所以选择在 $t_0 = 0$ 时开始观测。如果使用一个可变的开始时间,而且开始时间包含多于一个小数位(例如 $t_0 = 0.12345\text{s}$),必须使用插值来根据航迹数据生成笛卡儿采样值。如果原始航迹包括4000个采样点,最终会得100个采样点来生成观测值集合。

假设降采样后的球坐标真实数据包含在可变极坐标真实数据(维数为 $P \times N$,其中 P 是观测矢量的维数,而 N 是时间采样点的数目)中,用列表18-1的Matlab程序片段将噪声添加到真实航迹采样值中(Stream是一个标志符,为最新版本的randn函数中的随机数流)。

列表18-1 加入随机观测噪声

```
Observations = PolarTruth + sqrtm(R) * randn(Stream,P,N);
```

为每次蒙特卡洛运行生成一组独立的观测值。每次蒙特卡罗运行最后的观测值集合包含 N 个样本。这一过程可以离线完成,且与跟踪性能计算无关。

18.3.3 滤波器初始化

所有序贯卡尔曼跟踪滤波器都需要计算状态矢量和协方差矩阵,分别为 $\hat{\boldsymbol{x}}_{0|0}$ 和 $\boldsymbol{P}_{0|0}^{xx}$。假设要访问传感器数据流(如上所述,可以是仿真的,也可以是真实的),最简单的方法是利用开始的少数传感器观测数据,来创建笛卡儿状态矢量的初始估计,进而生成所需的初始值。

如果用一个完整的含噪球极坐标6维状态矢量作为观测矢量,可以通过简单的对含噪球坐标状态矢量数据的一些样本取平均,然后用式(18-195)定义的极坐标—笛卡儿坐标变换,来生成笛卡儿状态矢量的初始值。但是在数据矢量中,最多只能有6个球坐标分量中的4个。

由于不可能从观测值中得球坐标状态矢量的所有分量,对那些无法直接由传感器数据中获取的球坐标分量,需要设置默认值。最简单的实现方式是为球坐标状态矢量和它们对应的标准偏差的所有元素生成一组额定值。一般来说,所有的传感器都提供方位及方位的不确定性信息,假定总是可以从传感器获得这些数据,不用设置额定值。已经创建一组额定的球坐标初始值,用在所有的目标跟踪滤波器当中进行性能分析,这些都列在表18-1中。在球坐标初始协方差矩阵

中,将所有非对角项设置为零。

表 18 - 1 初始状态矢量和协方差元素的额定值

	球坐标矢量元素额定值	标准偏差额定值
距离 /n mile	10	0.1
距离变化率 /kn	-600	5
方位 /rad	可由传感器获取	可由传感器获取
方位变化率 /(rad/s)	0	10^{-4}
高程 /rad	0	10^{-2}
高程变化率 /(rad/s)	0	10^{-4}

当得一个或多个球坐标元素观测数据时,可以根据连续的样本数据计算初始值。根据位置数据(R,θ,ϕ),通过对连续数据值取平均来生成一个初始值,例如可以对前两个数据样本取平均,得距离的初值,即

$$R_0 = \frac{1}{2}(R_1 + R_2) \tag{18 - 108}$$

对于速率元素,可以用前3个数据点的2个,由式(2 - 44)来计算中心差分公式,例如可以计算距离变化率为

$$\dot{R}_0 = \frac{1}{2}(R_3 - R_1) \tag{18 - 109}$$

当处理异步多传感器数据时,必须进行一些改进,因为某些分量可能从一种传感器中获取,而在另一种传感器中得不到。

如果得了完整的球坐标状态矢量和协方差矩阵的初始值,就可以用附录18.B给出的公式进行简单的球坐标—笛卡儿坐标变换,求得笛卡儿状态矢量和协方差矩阵的初始值。要确保笛卡儿坐标系下初始状态矢量和协方差矩阵集合是给定的,并且对于所有蒙特卡洛运行和所有跟踪估计方法都是相同的。

对于球坐标恒速率滤波,这些笛卡儿的初始值必须变换回球坐标,生成集合$\{\boldsymbol{r}_{0|0},\boldsymbol{P}_{0|0}^{rr}\}$,并将其转换为初始化集合为$\{\hat{\boldsymbol{\rho}}_{0|0},\boldsymbol{P}_{0|0}^{\rho\rho}\}$的恒速度球坐标滤波器。注意,由$\boldsymbol{r}$到$\boldsymbol{\rho}$的变换由下式给出,即

$$\boldsymbol{\rho} = \begin{bmatrix} R \\ v_R \\ \theta \\ v_h \\ \phi \\ v_v \end{bmatrix} = \boldsymbol{g}(\boldsymbol{r}) = \begin{bmatrix} R \\ \dot{R} \\ \theta \\ \dot{\theta}R\cos\phi \\ \phi \\ \dot{\phi}R \end{bmatrix} \tag{18 - 110}$$

$$\boldsymbol{P}^{\rho\rho} = \hat{\boldsymbol{G}}\boldsymbol{P}^{rr}\boldsymbol{G}^{\mathrm{T}} \tag{18 - 111}$$

式中：$\hat{G}$ 为变换的雅可比矩阵 $\boldsymbol{g}(\boldsymbol{r})$。计算雅可比矩阵的偏导，得

$$\hat{G}=\begin{bmatrix}1 & 0 & 0 & 0 & 0 & 0\\ 0 & 1 & 0 & 0 & 0 & 0\\ 0 & 0 & 1 & 0 & 0 & 0\\ \dot{\theta}\cos\phi & 0 & 0 & R\cos\phi & -R\dot{\theta}\sin\phi & 0\\ 0 & 0 & 0 & 0 & 1 & 0\\ \phi & 0 & 0 & 0 & 0 & R\end{bmatrix} \tag{18-112}$$

18.4 各种估计方法的性能对比

18.4.1 用于性能分析的轨迹特征

为便于比较分析，用附录18.A列出的方法生成3个目标运动场景。图18－1至图18－3分别表示与3个场景其中一个相关的航迹。

第一个场景航迹如图18－1所示，是以恒定笛卡儿速度径向运动的目标。这个航迹中，目标以1n mile的恒定径向入站高度（在图18－1中没有显示高度）运动，并直接通过位于(0,0,0)的雷达传感器。第二个场景航迹如图18－2所示，是进行两次转弯的目标，第一次是90°2g的转弯，第二次是90°5g的转弯。第三个航迹是一个高度机动目标，如图18－3所示，目标首先经历一个7g的转弯后，紧接着一个6g的转弯，减速后进行5g的俯冲，平稳下来后又加速，最后是7g的转弯。所有这些场景都类似于Zollo和Ristic[5]所提出的场景。

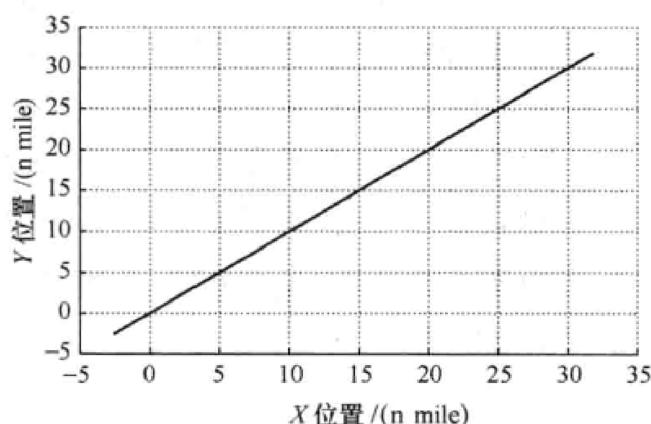


图18－1　仿真径向入站目标真实轨迹

表18－2给出了3个场景的具体初始特征值。对于用于性能分析的所有跟踪滤波器，将 q 的分量值设置为：笛卡儿跟踪器 $\{q_x=q_y=225; q_z=1.5\}$，球坐标恒速度跟踪器 $\{q_r=q_h=225; q_v=1.5\}$，球坐标恒速率跟踪器 $\{q_r=225; q_{\dot{\theta}}=q_{\dot{\phi}}=1.5\}$。$q$ 值取225还是1.5，取决于预期的最大转弯／俯冲加速度为7.5g还是0.61g。

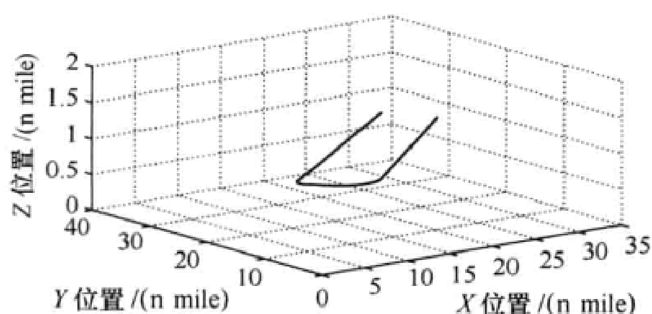


图 18 - 2　水平转弯目标真实轨迹仿真

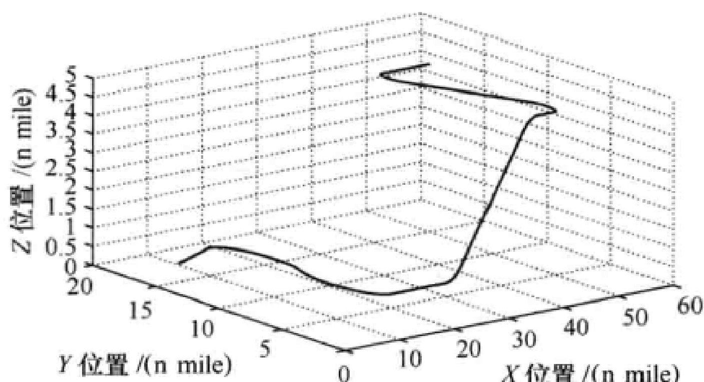


图 18 - 3　仿真基准测试目标真实轨迹

表 18 - 2　3 个仿真场景的初始特征值

	航迹		
	径向航迹	环形航迹	基准航迹
速度 /kn	350	700	700
初始位置 /n mile(x,y,z)	(31.82,31.82,1)	(30.78,30.78,1)	(50,20,3)
初始距离 /n mile	45	40	58.58
初始方位	45°	45°	70°
初始朝向	- 135°	- 135°	- 90°
初始姿势	0°	0°	0°

首先检查每个场景中速度特征的航迹，来演示如何使航迹符合任一种恒速度模型。图 18 - 4 至图 18 - 6 给出了 3 个场景在笛卡儿坐标系和球坐标系下的速度分量。根据图 18 - 4，径向入站目标的笛卡儿速度很显然是恒定的，而当轨迹在原点穿过传感器的区域时，距离和垂直速度的球坐标速度分量都有突然变化的迹象。因此，可以预测：远离原点区域时两个跟踪器模型之间的性能差异很小，但是当目标接近原点时球坐标跟踪器会遇到一定的困难。虽然在图中没有显示，但是，当目标经过原点时，由于突然变化，恒速率球坐标滤波器中的速率在很短的距离中显示出较大的变化。

比较图 18 - 5 中环形航迹的场景速度，可见速度在笛卡儿坐标和球坐标中

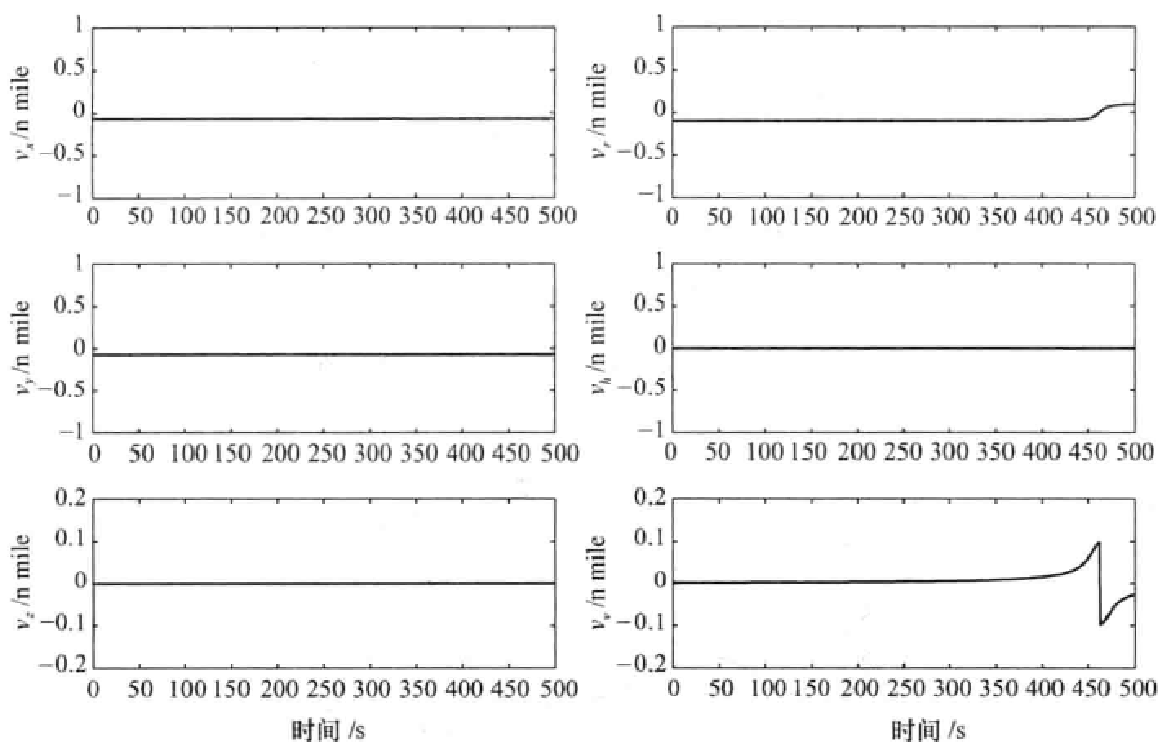


图 18 －4　径向入站航迹的笛卡儿速度和球坐标速度分量

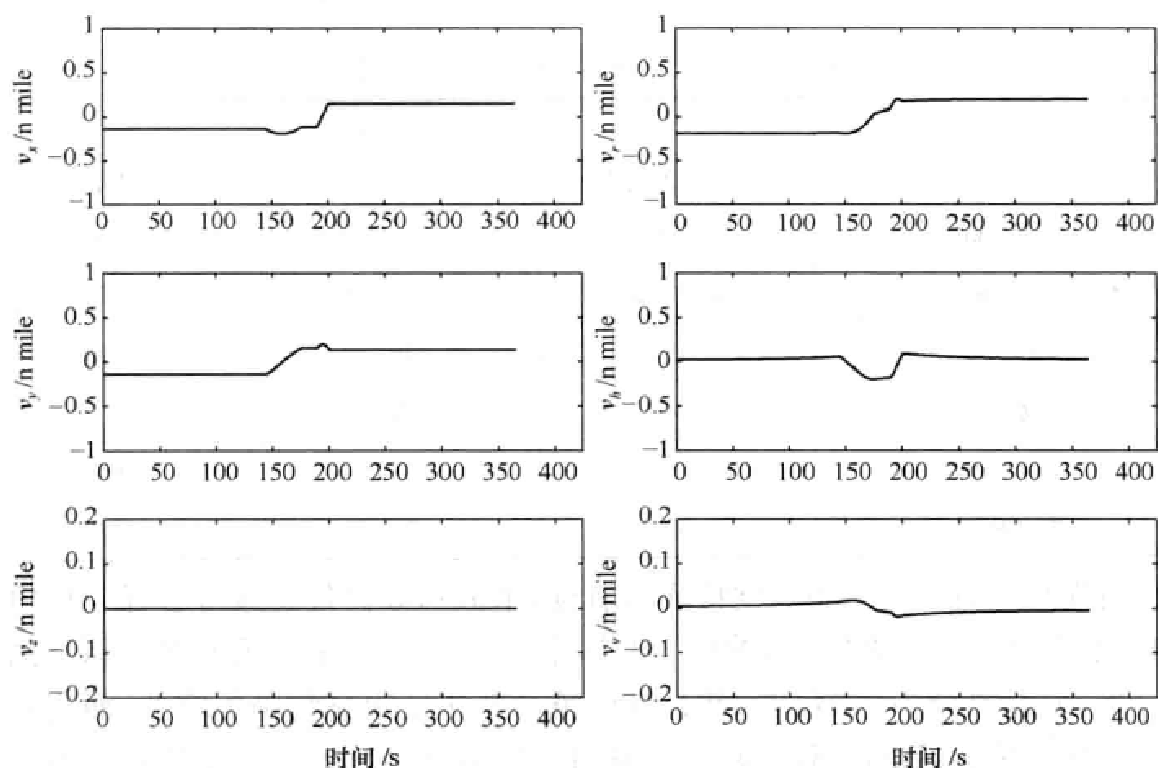


图 18 －5　环形航迹的笛卡儿速度和球坐标速度分量

显示相似的变化，所以跟踪性能的差异应该比较小。对于如图 18 －6 所示高度机动航迹，由于作为时间的函数，其速度上的变化相似，图中速度分量表明球坐标跟踪器的运行效果应该和笛卡儿跟踪器相似。

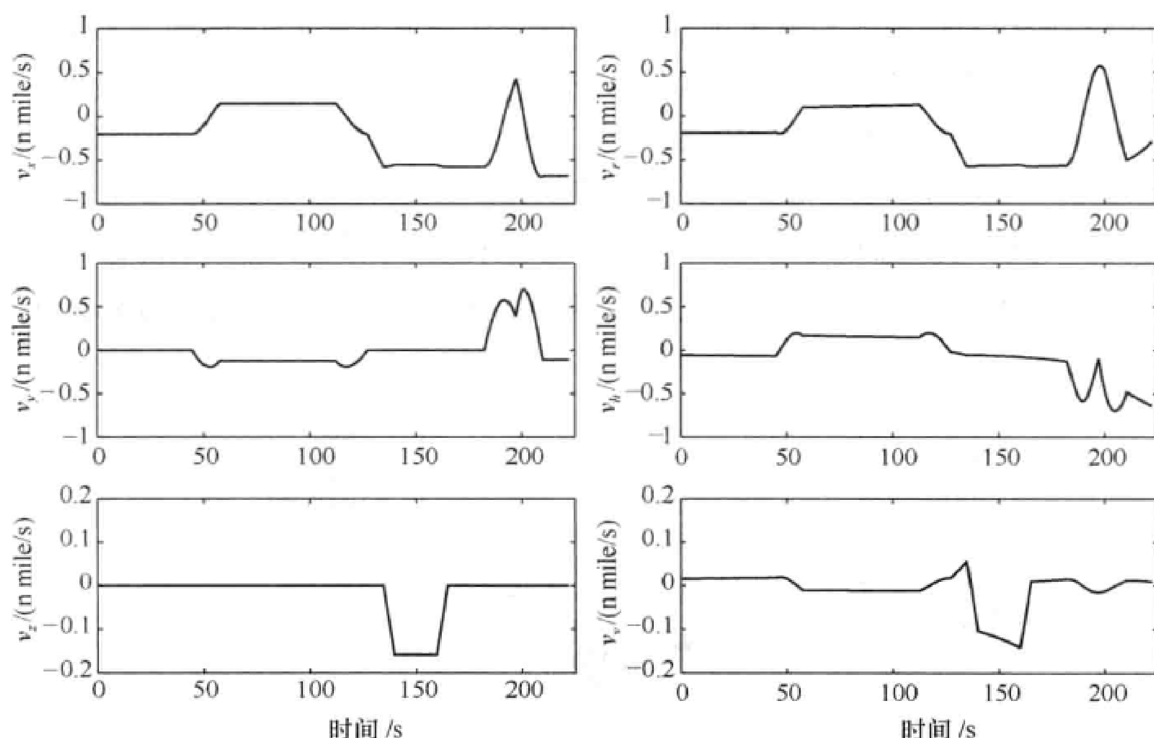


图 18 - 6　基准测试航迹的笛卡儿速度和球坐标速度分量

18.4.2　滤波器性能对比

在笛卡儿坐标和球坐标跟踪器中，采用相同的跟踪估计滤波器（EKF、UKF、SSKF 和 GHKF），为方便区分，分别用前缀 C - 和 S - 来区分笛卡儿坐标和球坐标跟踪算法。用 4 个笛卡儿坐标卡尔曼滤波器 C - EKF、C - UKF、C - SSKF 和 C - GHKF 来生成笛卡儿恒速度模型的跟踪轨迹，再用 4 个球坐标卡尔曼滤波器 S - EKF、S - UKF、S - SSKF 和 S - GHKF 来生成球坐标恒速度模型的跟踪轨迹。对于球坐标恒速率远距离估计模型，采用 S - LKF 滤波器。

对每种情况，进行 100 次的蒙特卡洛运行。对于笛卡儿跟踪算法，将笛卡儿坐标估计轨迹输出变换为球坐标；而对于球坐标跟踪器，将球坐标跟踪输出变换为笛卡儿坐标，所有的坐标变换都利用附录 18. B 中的算法进行计算。因此，对所有的跟踪器，生成笛卡儿坐标和球坐标跟踪估计。根据 100 次蒙特卡洛运行结果，在笛卡儿坐标系和球坐标系中计算均方根误差的位置估计。

18.4.2.1　径向入站跟踪航迹的跟踪性能比较

图 18 - 7 和图 18 - 8 给出分别用笛卡儿和球坐标系卡尔曼滤波器，径向入站目标第一次蒙特卡洛运行的三维估计轨迹和真实轨迹（灰色线条）。由这些估计跟踪图可以看出：当目标经过位于(0,0,0)的传感器时，球坐标跟踪器效果比笛卡儿跟踪器更好。除了 C - EKF，其它笛卡儿滤波器在原点附近区域表现不佳。

可以在笛卡儿跟踪估计方法的非线性观测方程(18 - 10)中找出笛卡儿跟

踪估计在原点附近发散现象出现的原因。尽管具体细节超出了本书的范畴,结果表明,式(18 - 10)中包括的反正切运算造成了发散。由于笛卡儿位置分量的大小接近于估计位置不确定性的大小,任意两个位置分量比值的反正切概率密度函数近似于均匀分布。因此方位和高程的预测都由越来越大的标准偏差控制,导致式(18 - 26)中的新息矢量$(z_n - \hat{z}_{n|n-1})$变得不稳定,进而导致了滤波器发散。然而,和预期一样,在远离原点的区域,笛卡儿滤波器比球坐标系滤波器效果更佳。

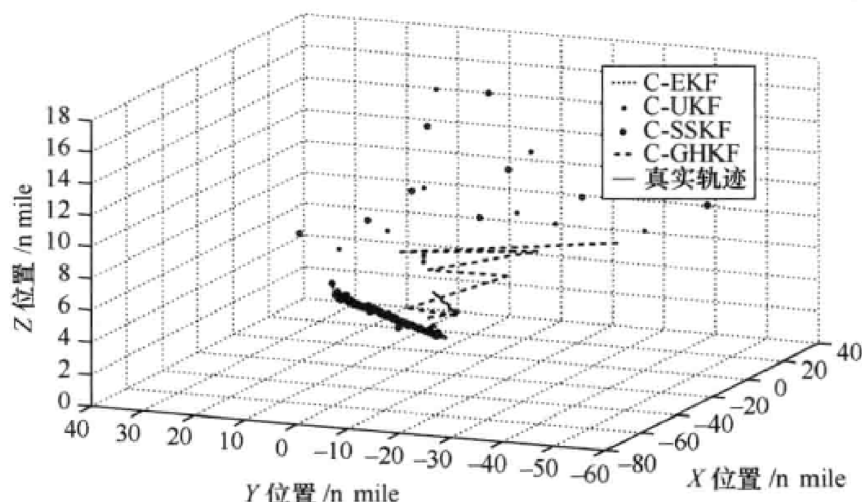


图 18 - 7 用笛卡儿跟踪方法计算径向入站目标航迹的估计轨迹

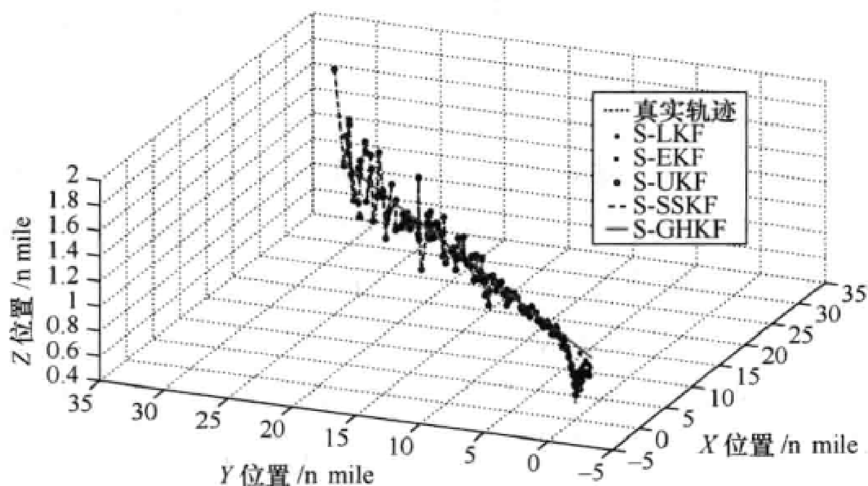


图 18 - 8 用球坐标跟踪方法计算径向入站目标航迹的估计轨迹

对于径向入站目标航迹,检查图 18 - 9 中笛卡儿坐标系和球坐标系跟踪算法的 RMS 笛卡儿位置误差对比,以及图 18 - 10 中两类跟踪算法的 RMS 球坐标系位置误差对比。在这些图中,设 $q_x = q_y = q_r = q_h = 225$,且所有其它 q 值均为 1.5。很明显,在估算笛卡儿轨迹时,所有的跟踪器运行得都很好,除了在大约 410s 目标穿过雷达区域时。在这一点之后,球坐标系跟踪器表现出更好的性能,其中 S - GHKF 是最好的。在近距离跟踪区域,球坐标跟踪器在高程上表现出很大的 RMS 误差,但对所有其它笛卡儿和球坐标变量给出了额定的 RMS 误差。

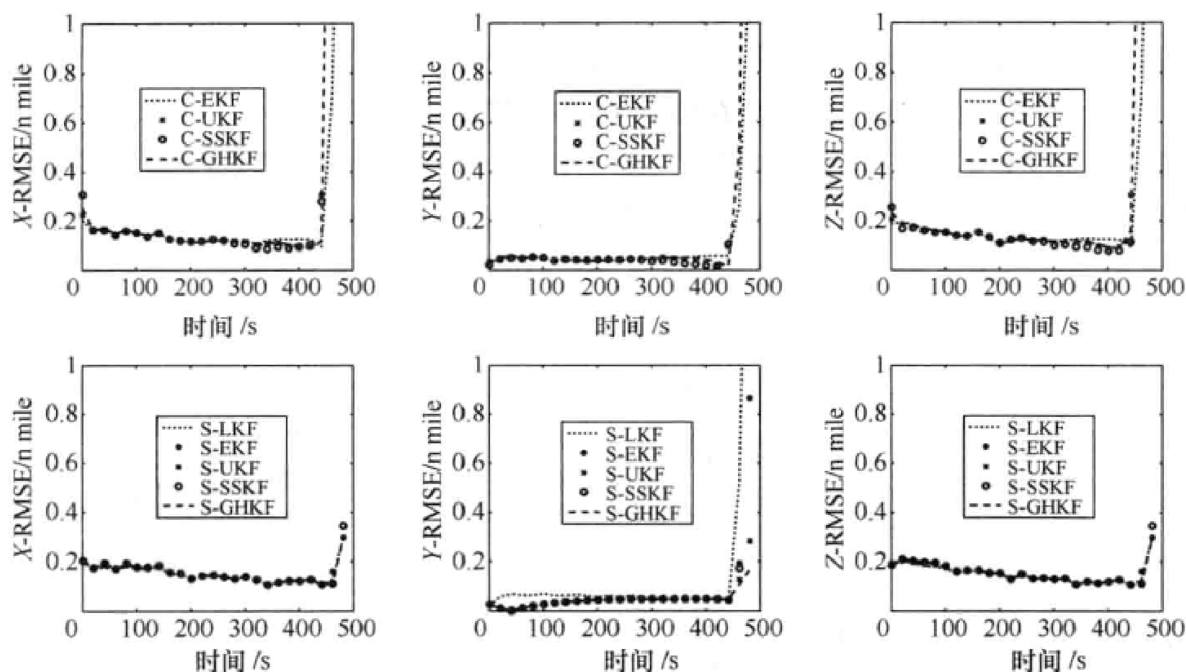


图 18 - 9　用笛卡儿坐标系和球坐标系跟踪算法计算径向入站目标航迹的 RMS 笛卡儿坐标位置误差

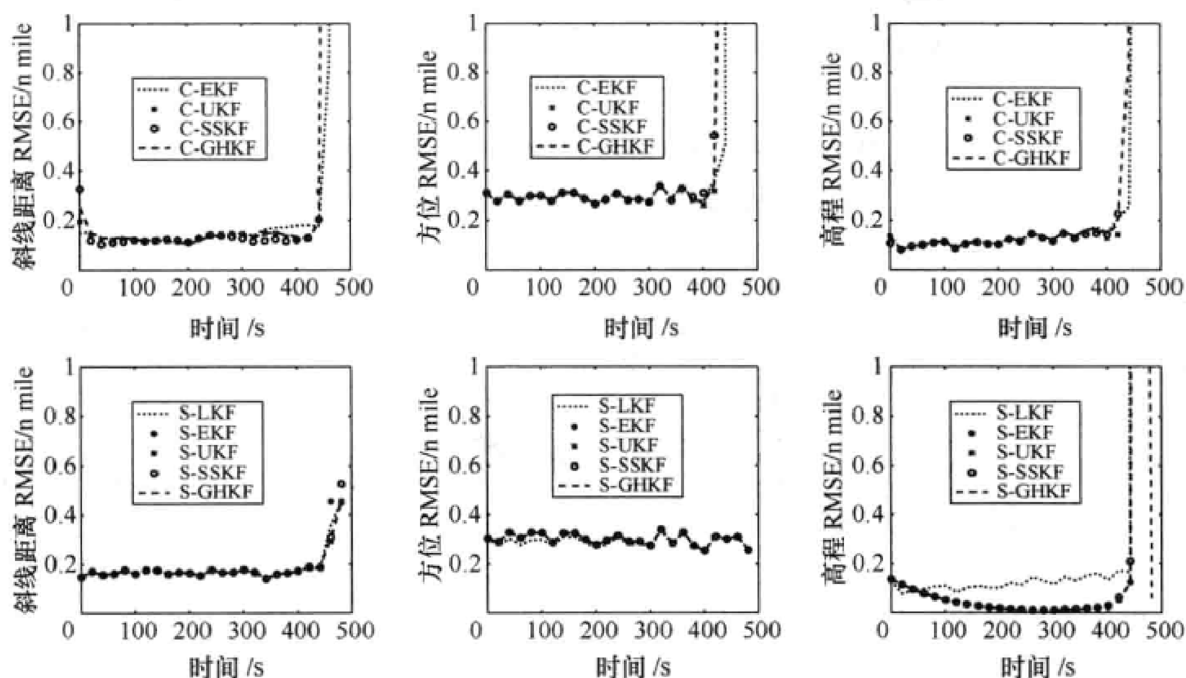


图 18 - 10　用笛卡儿坐标系和球坐标系跟踪算法计算径向入站目标航迹的 RMS 球坐标位置误差

现在检验径向入站航迹中 q 值变化对 RMS 位置误差性能的影响。要注意，在图 18 - 9 和图 18 - 10 中，笛卡儿滤波器族与球坐标滤波器之间的性能差别很小。为更加简单地检验不同 q 值对性能的影响，这里仅比较当 q 变化时，笛卡儿和球坐标系 EKF 算法的性能结果。尽管本书这里没有给出，但根据对 q_z、$q_{angular}$ 和 q_v 值

变化的初步研究结果，把它们设成以下三个值当中的一个：1.5、36 和 225。从分析中很显然可以看出，对于所有航迹最好的选择是 1.5。

由于径向入站航迹中包含在整个轨道保持恒定速度的目标，有读者会认为，RMS 位置误差将会与 q 的值成正比，即当 q 增加时，RMS 误差随之增加。然而，仔细观察图 18 - 11 和图 18 - 12，可看到：当 q 水平值为 36 时，笛卡儿跟踪估计性

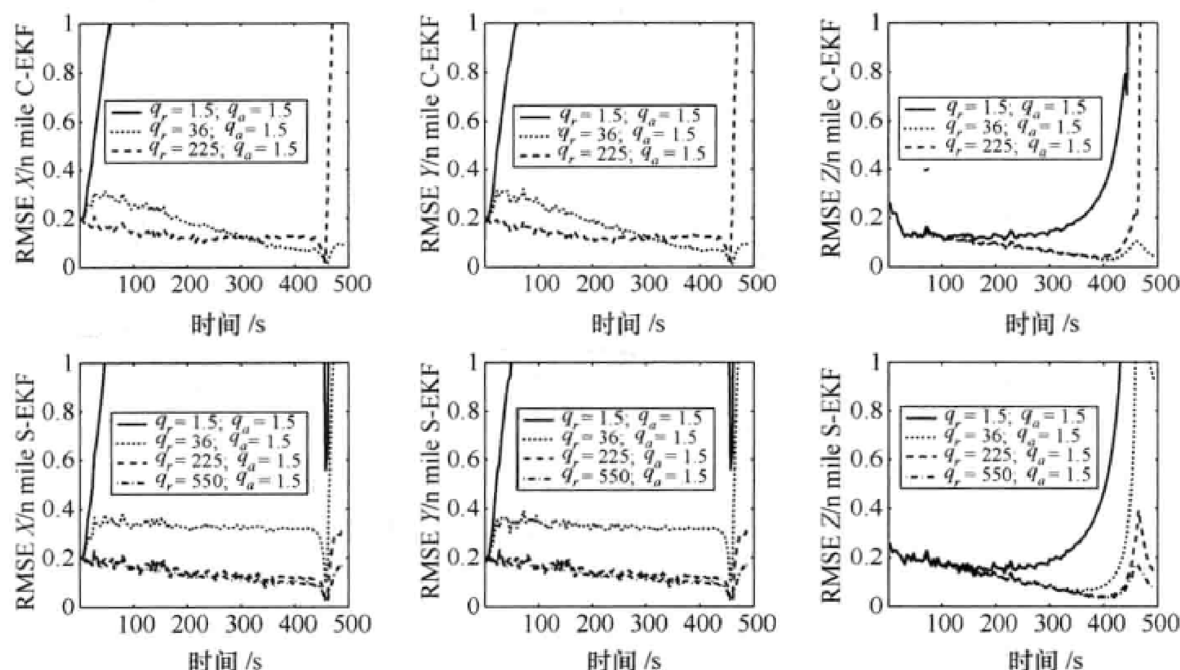


图 18 - 11　笛卡儿和球坐标系 EKF 滤波器中不同 q 值下，径向入站航迹 RMS 笛卡儿位置误差的比较

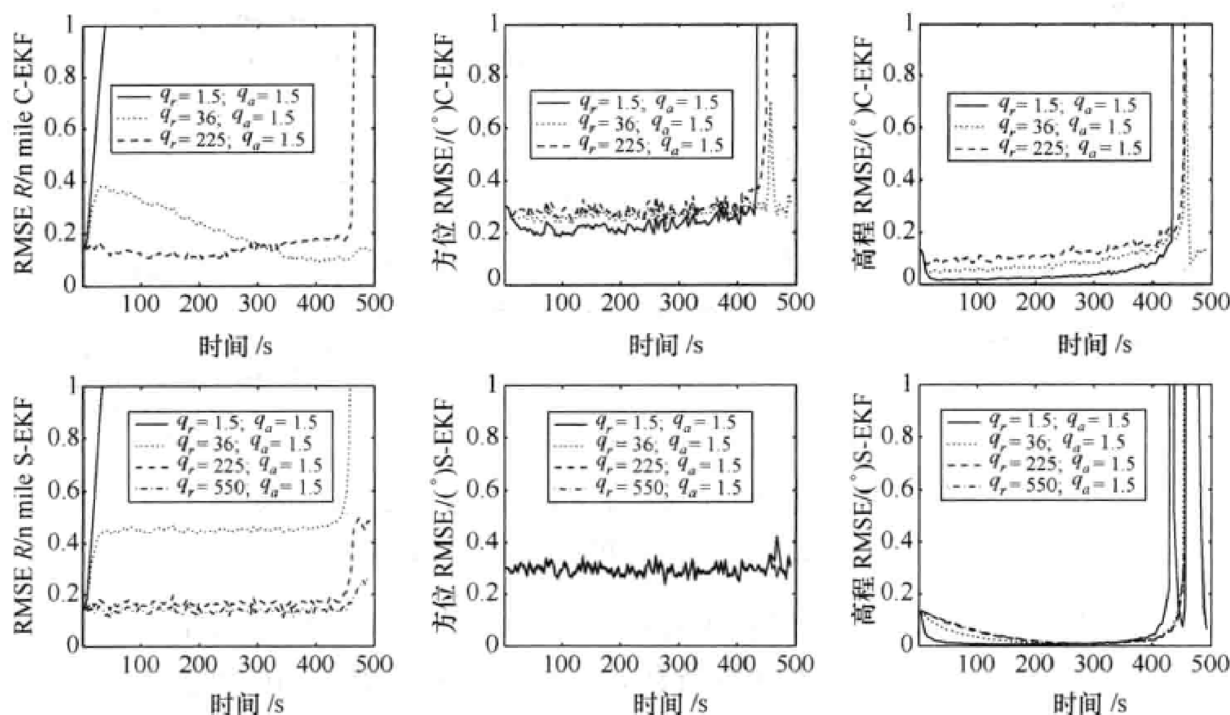


图 18 - 12　笛卡儿和球坐标 EKF 滤波器中不同 q 值下，径向入站航迹 RMS 球坐标位置误差的比较

能最佳;而对于球坐标跟踪器来说最好的选择是550。注意,对于笛卡儿滤波器算法,q_r 和 q_a 分别表示水平和垂直的 q 值。由图可见,所有的滤波器在原点附近高程估计的RMS误差都很大,球坐标系滤波器的性能略好。比较结果为,当 q_r = 550时的S – EKF的性能最好。

18.4.2.2 环形目标航迹跟踪性能比较

在环形航迹场景中,图18 – 13和图18 – 14分别给出了笛卡儿跟踪算法和球坐标跟踪算法中,沿真实轨迹输出跟踪轨迹的例子。对图例进行观察,可见所有的追踪器除了在高 g 转弯的地方稍有延迟,其余都表现很好,笛卡儿跟踪器在远离转弯区域时的波动稍小。

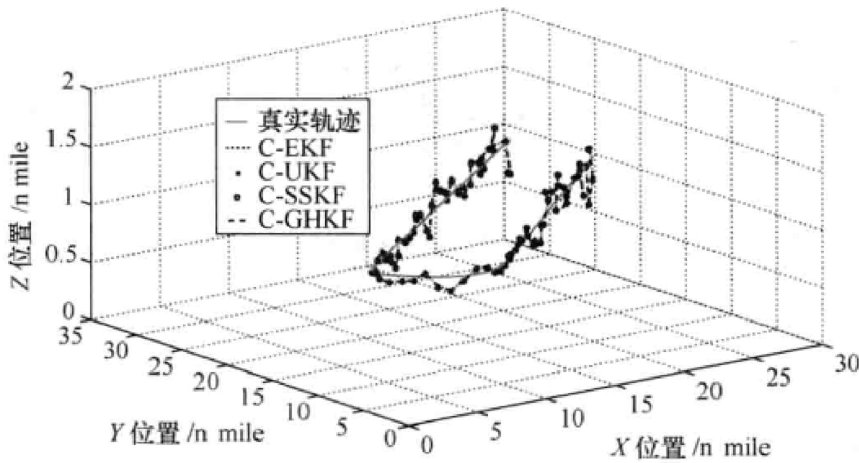


图18 – 13　笛卡儿跟踪方法在环形目标航迹估计中的跟踪结果

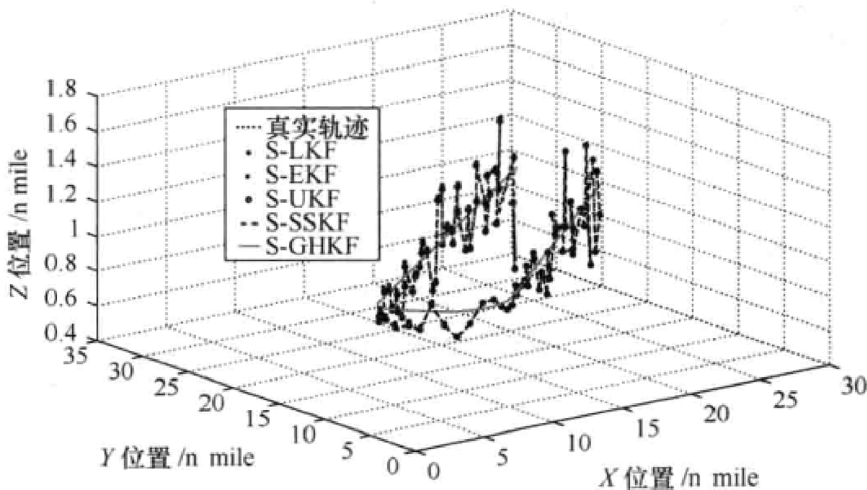


图18 – 14　所有球坐标跟踪方法在环形目标航迹中的估计轨迹

RMS位置误差如图18 – 15和图18 – 16所示,表明所有的跟踪估计算法具有相似的性能,笛卡儿跟踪算法的整体性能稍好。在笛卡儿跟踪器中,C – SSKF是性能最差的跟踪估计算法;而在球坐标跟踪器中,S – LKF是性能最好的。在这两个图中,设 $q_x = q_y = q_r = q_h = 225$,并将所有 q 值都设为1.5。

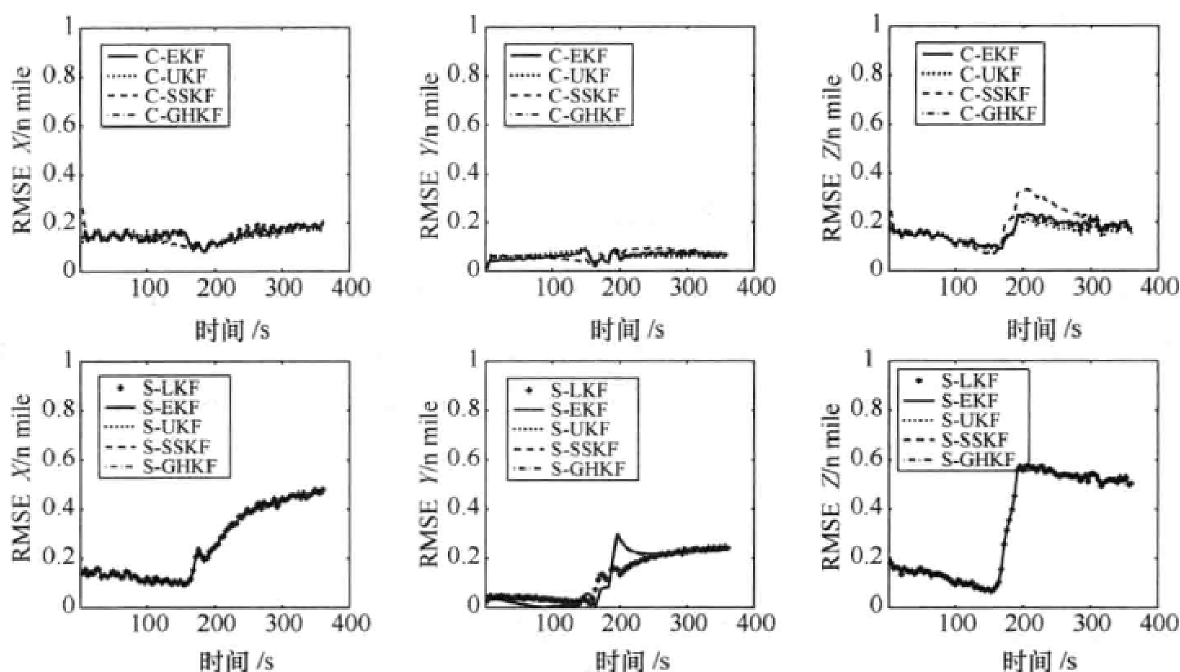


图 18－15　笛卡儿跟踪算法在环形目标航迹中的 RMS 笛卡儿和球坐标位置误差

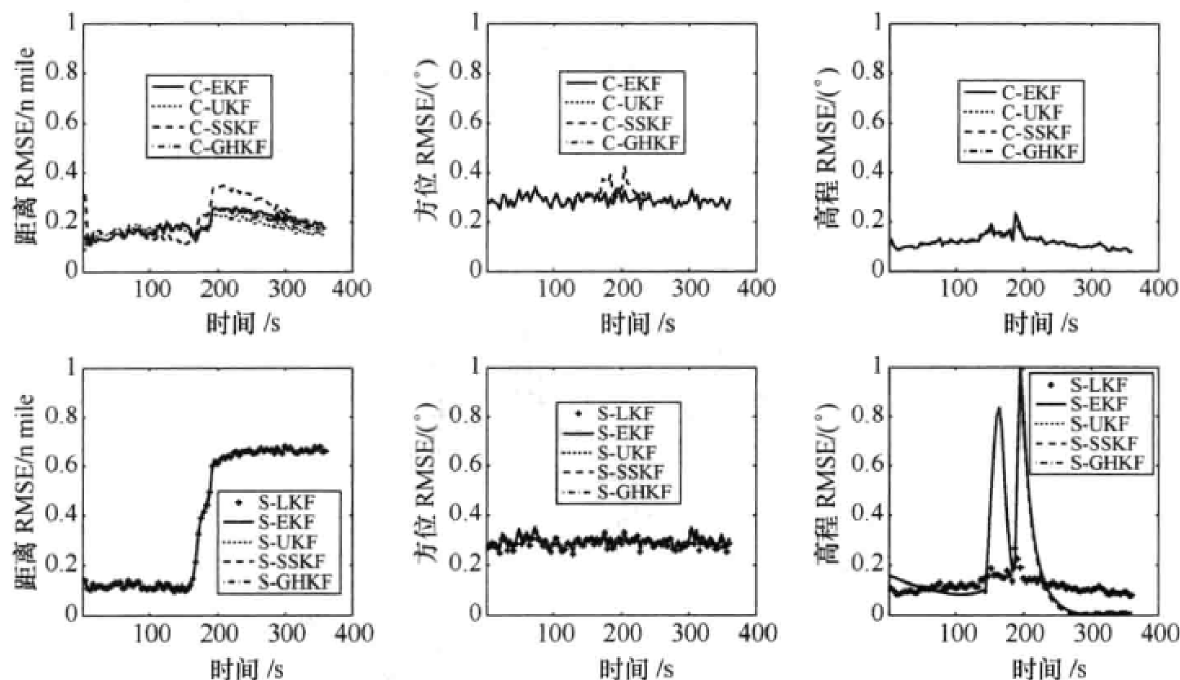


图 18－16　球坐标跟踪算法在环形目标航迹中的 RMS 笛卡儿和球坐标位置误差

为检验不同 q 值的效果，只采用笛卡儿和球坐标 EKF 来考查性能，其中 q_z、q_{angular} 和 q_v 值设为 1.5。图 18－17 和图 18－18 分别给出了 RMS 笛卡儿和球坐标位置误差的结果。结果表明，对于笛卡儿和球坐标 EKF，当水平 q 值分别设为 225 和 550 时，跟踪结果最好。

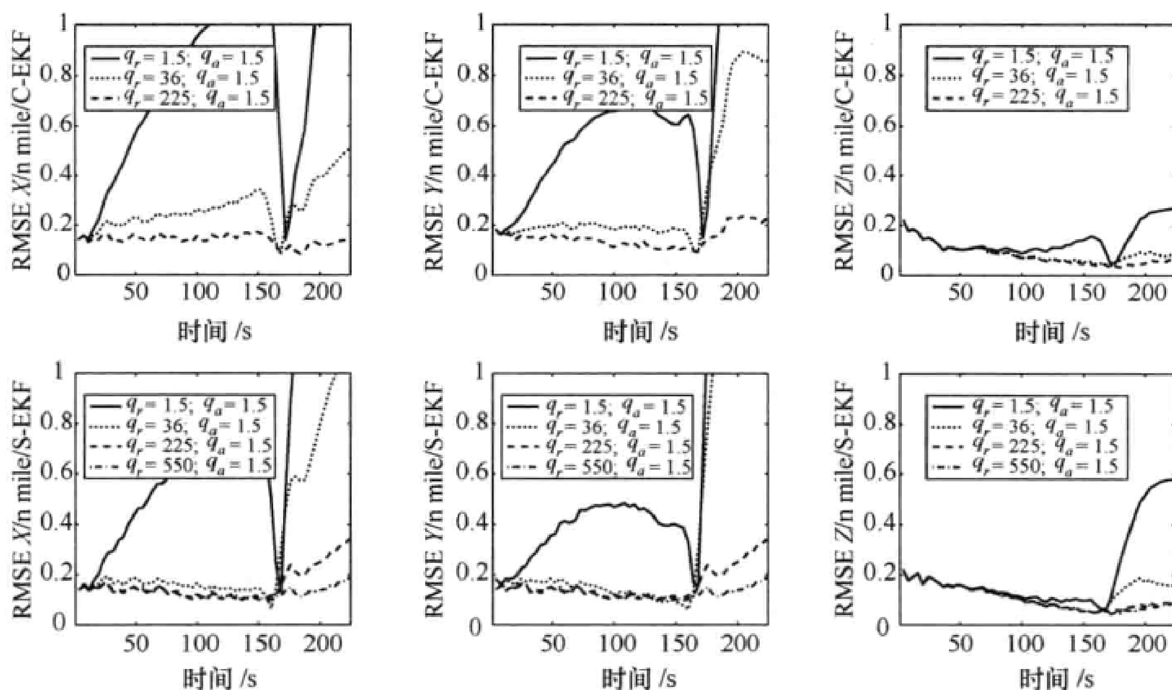


图 18 - 17　在环形航迹中,作为 q 的函数,RMS 笛卡儿位置误差性能比较

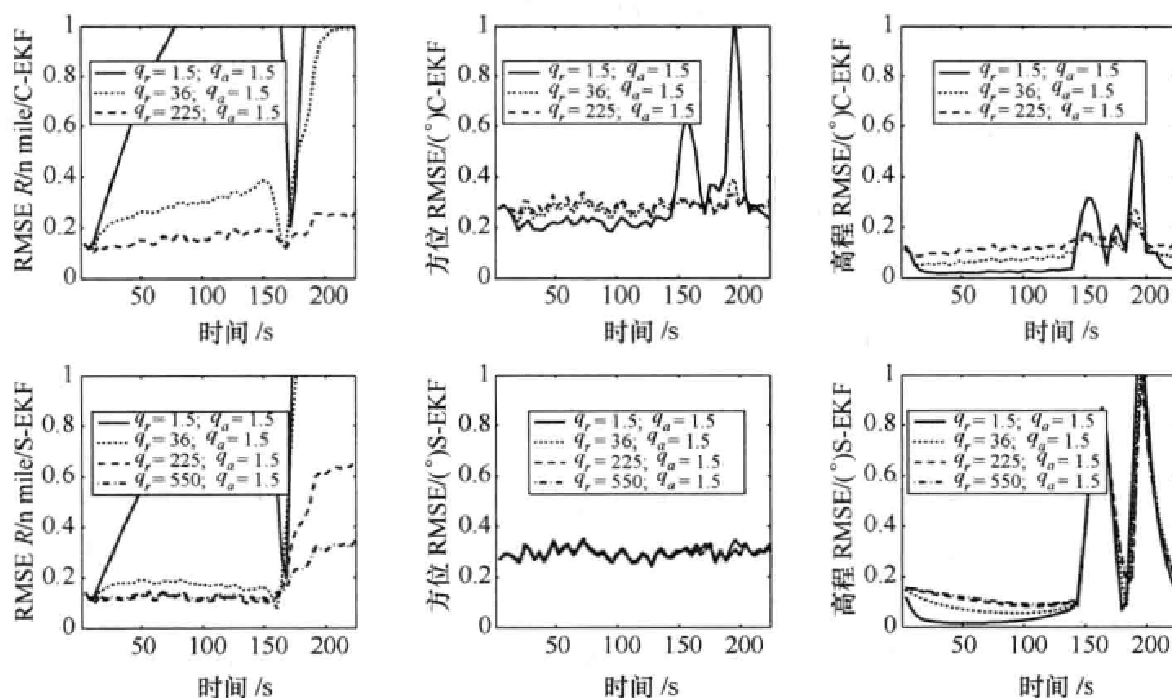


图 18 - 18　在环形航迹中,作为 q 的函数,RMS 球坐标位置误差性能比较

18.4.2.3　高度机动目标航迹基准测试的跟踪性能比较

图 18 - 19 和图 18 - 20 分别为采用笛卡儿和球坐标跟踪算法,对最后一个航迹的跟踪估计,其中有一些高 g 机动飞行,还有减速和加速部分。通过比较两图可见,笛卡儿系和球坐标系的所有的跟踪算法,除了在图表末端的加速部分,其余效果都很好。笛卡儿跟踪器在加速之后最后一个转弯处越过了目标;而球坐标

跟踪器没有这个最后转弯的问题。从这些图中，很难确定跟踪性能最好的跟踪算法。

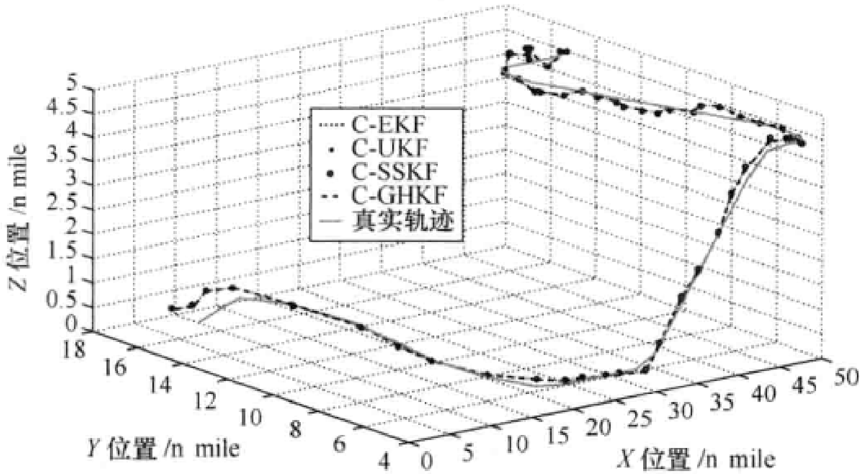


图 18 - 19　所有笛卡儿跟踪方法在基准目标航迹测试中的估计轨迹

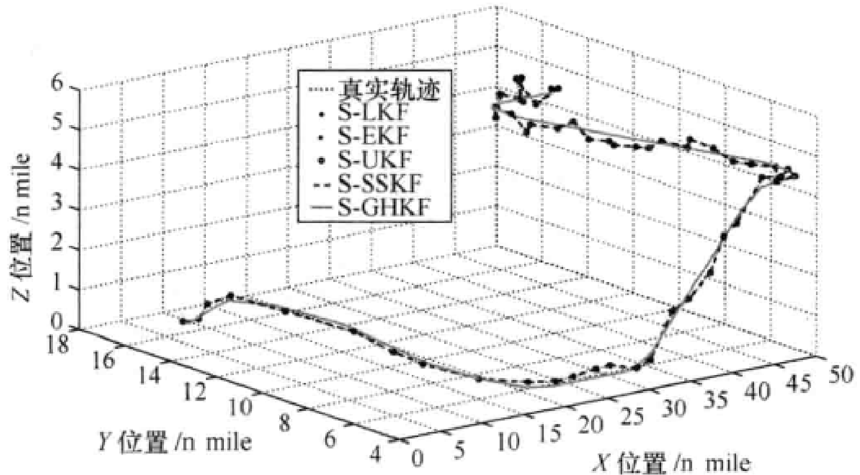


图 18 - 20　所有球坐标跟踪方法在基准目标航迹测试中的估计轨迹

图 18 - 21 和图 18 - 22 的 RMS 位置误差分别对应笛卡儿和球坐标 RMS 位置误差。比较可知，对于高度机动目标，与笛卡儿坐标系跟踪算法相比，球坐标系跟踪算法的 RMS 笛卡儿和距离位置误差稍低。此外，球坐标系跟踪估计算法的方位跟踪性能比较好，但在高程跟踪上性能最差。综上所述，在机动目标跟踪过程中，球坐标跟踪滤波器算法得的跟踪估计与笛卡儿跟踪估计算法效果相似，稍优于笛卡儿算法。还应指出的是，在所有滤波器中，恒定球速率滤波器的性能最好。对这种目标距离不小于 20n mile 的远距离跟踪场景，这并不奇怪。在这两个图中，设 $q_x = q_y = q_r = q_h = 225$，而其它的 q 值设为 1.5。

为研究不同 q 值对航迹的影响，只检验笛卡儿和球坐标 EKF，其中设 q_z、q_{angular} 和 q_v 为 1.5。图 18 - 23 和图 18 - 24 中分别给出了 RMS 笛卡儿和球坐标位置误差的结果，结果表明，对于笛卡儿和球坐标 EKF，当水平 q 值分别设为 225 和 550 时，跟踪结果最好。

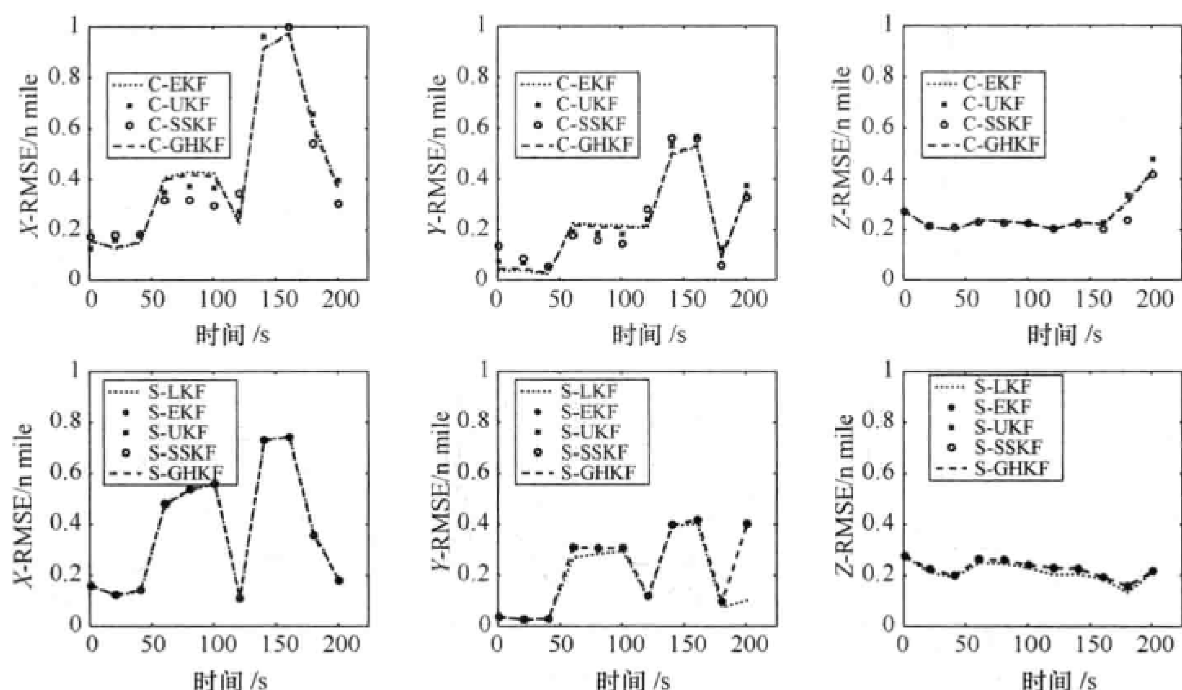


图 18－21　笛卡儿跟踪算法在基准目标航迹中的 RMS 笛卡儿和球坐标位置误差

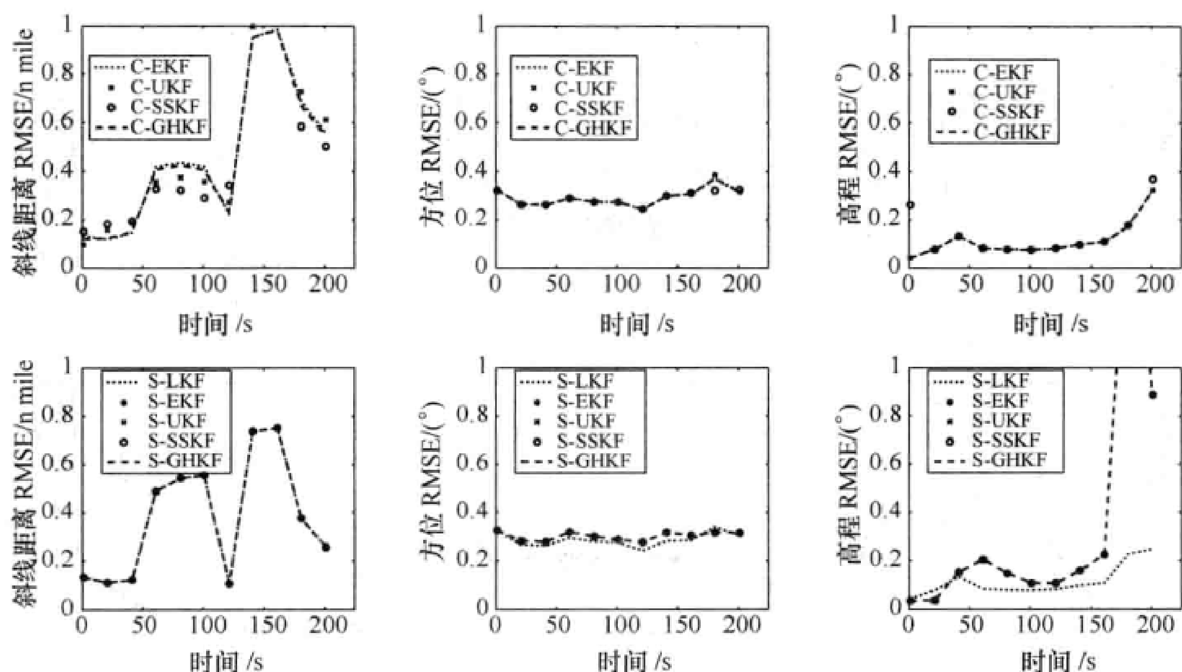


图 18－22　球坐标跟踪算法在基准目标航迹中的 RMS 笛卡儿和球坐标位置误差

18.4.2.4　总结

对于3个不同的目标轨迹，机动加速度逐个递增。考虑跟踪算法的 RMS 位置误差性能，可以得出结论：如果能正确选择 q 的分量，那么在机动目标跟踪中，新的球坐标恒速度动态模型优于笛卡儿坐标恒速度模型。而且，当目标接近距离 R 很小的区域时，新的跟踪器表现出性能的改善，特别是对于方位和高程的估计。

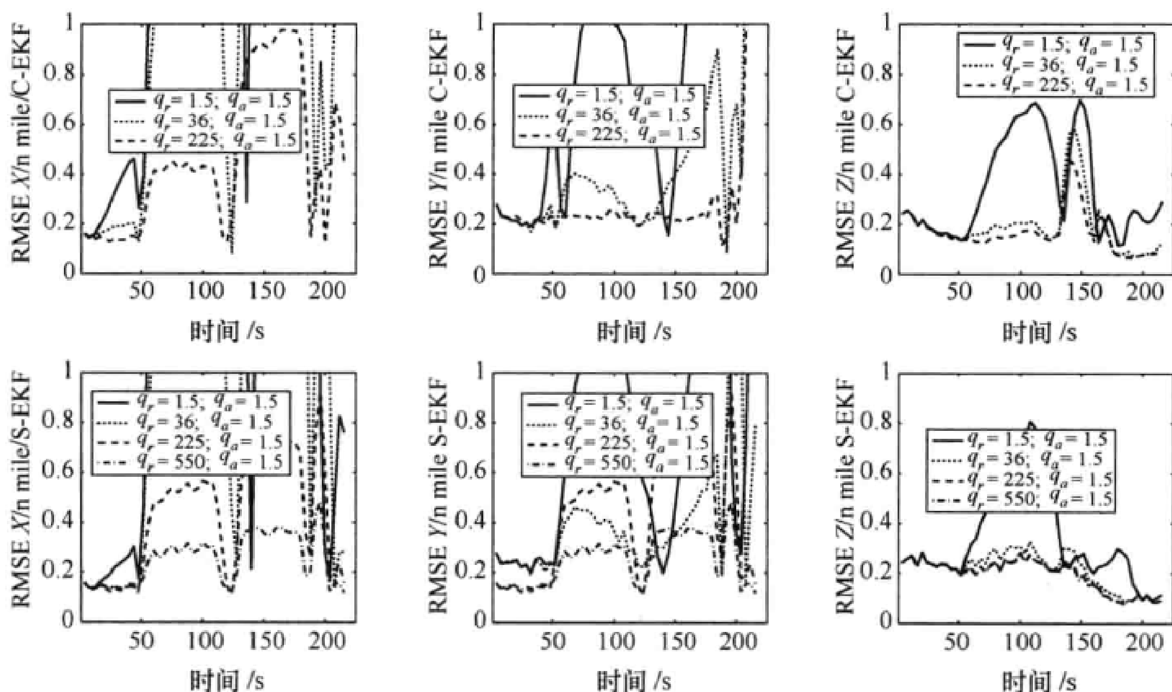


图 18 - 23　在基准航迹中，作为 q 的函数，RMS 笛卡儿位置误差性能比较

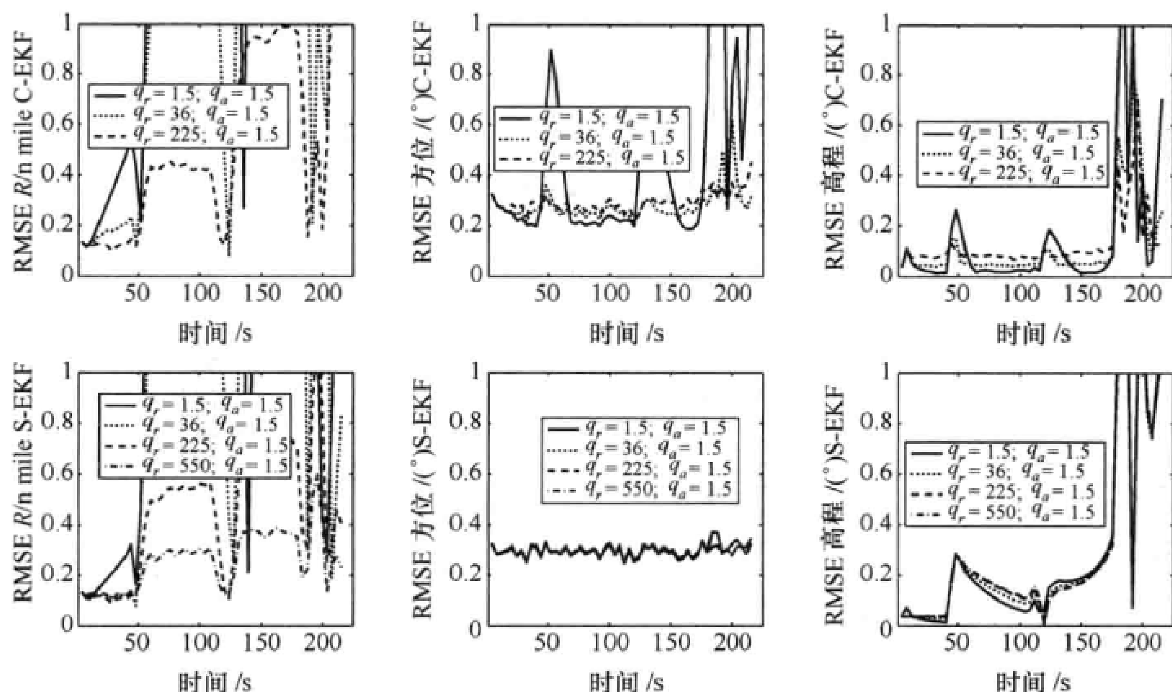


图 18 - 24　在基准航迹中，作为 q 的函数，RMS 球坐标位置误差性能比较

18.5　观测数据与展望

这个案例研究被认为是一个新的恒定球速度动态目标运动模型，可以将合成卡尔曼滤波器跟踪算法应用到这个模型中。通过对 3 个仿真航迹的性能分析，目标机动水平逐渐增加，可以看到：当目标为机动的且目标距离接近零时，球坐

标跟踪估计滤波器的性能更好。结果还表明，除了S－LKF，所有的球坐标跟踪估计算法的性能相同。后面的推导可得出结论，应将S－EKF应用于恒定球速度跟踪，因为其计算时间最短。

在这个案例中，由于空间和时间的限制，本书未能充分探究新的可能的动态模型。以上研究可能包括以下有意义的拓展：

（1）进一步研究混合高q模型和低q模型的球坐标IMM滤波器的可能应用领域；

（2）进一步研究距离变化率观测值对球坐标滤波器和笛卡儿滤波器性能的影响。

附录18.A　三维恒定角速率偏转运动动力学

本节提出一种生成航迹的方法，航迹中包括水平和垂直方向恒定转弯速率的转弯。通过该方法，可生成包含水平转弯、垂直俯冲和盘旋的运动对象航迹。

假设目标在t时刻以水平方向$\vartheta(t)$和垂直姿势$\gamma(t)$运动。采用图18－25中的向东－向北－向上笛卡儿坐标系，这里东北向的飞机是水平飞行的飞机。

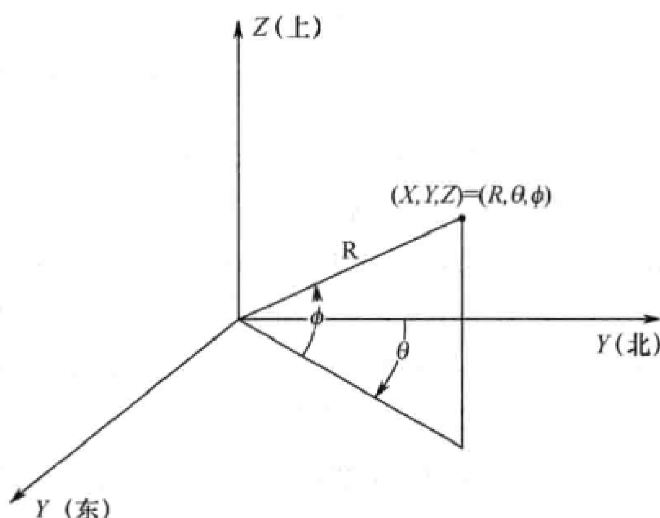


图18－25　东－北－上笛卡儿坐标系

18.A.1　恒定角速率偏转运动速度分量

可将恒定速度v在东－北－上笛卡儿坐标系中的笛卡儿速度分量表示为

$$\boldsymbol{v}(t)=\begin{bmatrix}v_x(t)\\v_y(t)\\v_z(t)\end{bmatrix}=\begin{bmatrix}v\sin\vartheta(t)\cos\gamma(t)\\v\cos\vartheta(t)\cos\gamma(t)\\v\sin\gamma(t)\end{bmatrix}\qquad(18-113)$$

式中：ϑ和γ分别为前进方向$-180^\circ\leqslant\vartheta<180^\circ$和姿势$-90^\circ\leqslant\gamma\leqslant90^\circ$。

对于一个恒速转动，参见图 18 - 26，可以有

$$\vartheta(t) = \vartheta(t_0) + \Omega(t - t_0) \triangleq \vartheta_0 + \Omega T \tag{18 - 114}$$

$$\gamma(t) = \gamma(t_0) + \Psi(t - t_0) \triangleq \gamma_0 + \Psi T \tag{18 - 115}$$

这里 $T \triangleq t - t_0$ 且 $\vartheta(t_0) \to \vartheta_0$，$\gamma(t_0) \to \gamma_0$。其中，$\Omega$ 和 Ψ 分别代表水平和垂直方向的恒定角速率。需要注意：当相对运动方向向右转弯时，Ω 为正向；向上转弯时，Ψ 为正向。将式(18 - 114) 和式(18 - 115) 代入式(18 - 113)，得

$$\begin{bmatrix} v_x(t) \\ v_y(t) \\ v_z(t) \end{bmatrix} = \begin{bmatrix} v\sin(\vartheta_0 + \Omega T)\cos(\gamma_0 + \Psi T) \\ v\cos(\vartheta_0 + \Omega T)\cos(\gamma_0 + \Psi T) \\ v\sin(\gamma_0 + \Psi T) \end{bmatrix} \tag{18 - 116}$$

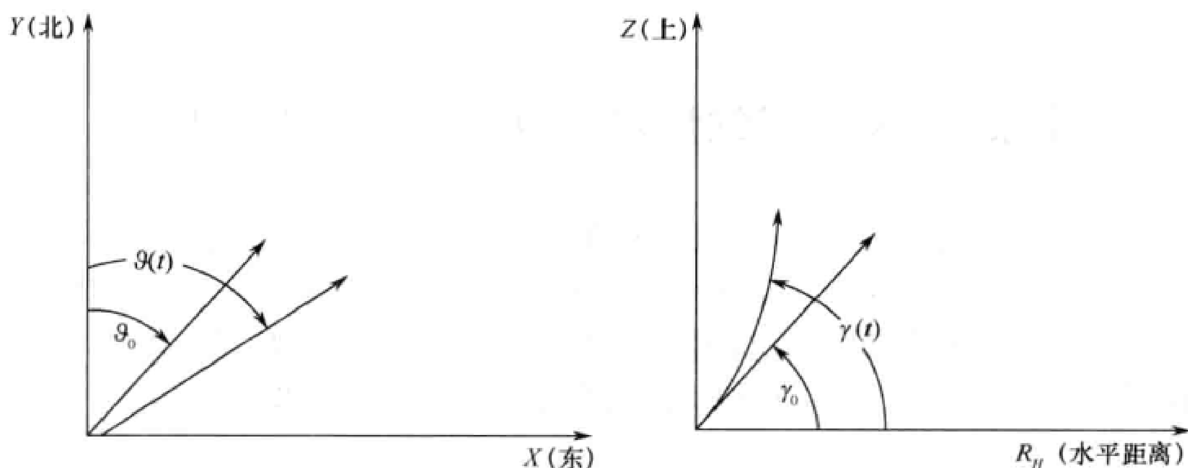


图 18 - 26　在水平和垂直方向以恒定角速率飞行的飞机

用角度和的正弦和余弦三角等式对 $v_x(t)$ 进行展开，可以将速度的分量写为

$$\begin{aligned} v_x(t) &= v\sin(\vartheta_0 + \Omega T)\cos(\gamma_0 + \Psi T) \\ &= v[\sin\vartheta_0\cos\Omega T + \cos\vartheta_0\sin\Omega T][\cos\gamma_0\cos\Psi T - \sin\gamma_0\sin\Psi T] \\ &= v\sin\vartheta_0\cos\gamma_0\cos\Omega T\cos\Psi T - v\sin\vartheta_0\sin\gamma_0\cos\Omega T\sin\Psi T + \\ &\quad v\cos\vartheta_0\cos\gamma_0\sin\Omega T\cos\Psi T - v\cos\vartheta_0\sin\gamma_0\sin\Omega T\sin\Psi T \end{aligned} \tag{18 - 117}$$

用式(18 - 113) 中分量 $\boldsymbol{v}$ 的定义，将上式变为

$$\begin{aligned} v_x(t) = {} & v_x(t_0)\cos\Omega T\cos\Psi T + v_y(t_0)\sin\Omega T\cos\Psi T - \\ & v_z(t_0)[\cos\vartheta_0\sin\Omega T\sin\Psi T + \\ & \sin\vartheta_0\cos\Omega T\sin\Psi T] \end{aligned} \tag{18 - 118}$$

用乘积—求和三角恒等式[9]，将上式改写为

$$\begin{aligned} v_x(t) = {} & \frac{1}{2}\{v_x(t_0)\cos\xi_1 T + v_y(t_0)\sin\xi_1 T - \\ & v_z(t_0)[\cos\vartheta_0\cos\xi_1 T - \sin\vartheta_0\sin\xi_1 T] + \\ & v_x(t_0)\cos\xi_2 T + v_y(t_0)\sin\xi_2 T + \\ & v_z(t_0)[\cos\vartheta_0\cos\xi_2 T - \sin\vartheta_0\sin\xi_2 T]\} \end{aligned} \tag{18 - 119}$$

$$\xi_1 \triangleq \Omega - \Psi \tag{18 - 120}$$

$$\xi_2 \triangleq \Omega + \Psi \tag{18-121}$$

类似地，得

$$\begin{aligned} v_y(t) = \frac{1}{2}\{ & -v_x(t_0)\sin\xi_1 T + v_y(t_0)\cos\xi_1 T + \\ & v_z(t_0)[\cos\vartheta_0 \sin\xi_1 T + \sin\vartheta_0 \cos\xi_1 T] - \\ & v_x(t_0)\sin\xi_2 T + v_y(t_0)\cos\xi_2 T - \\ & v_z(t_0)[\cos\vartheta_0 \sin\xi_2 T + \sin\vartheta_0 \cos\xi_2 T]\} \end{aligned} \tag{18-122}$$

计算 $v_z(t)$，有

$$\begin{aligned} v_z(t) &= v\sin(\gamma_0 + \Psi T) \\ &= v\sin\gamma_0 \cos\Psi T + v\cos\gamma_0 \sin\Psi T \\ &= v_z(t_0)\cos\Psi T + v\cos\gamma_0 \sin\Psi T \end{aligned} \tag{18-123}$$

令

$$\begin{aligned} v_h(t_0) &= \sqrt{v_x^2(t_0) + v_y^2(t_0)} \\ &= \sqrt{v^2\sin^2\vartheta_0\cos^2\gamma_0 + v^2\cos^2\vartheta_0\cos^2\gamma_0} \\ &= v\cos\gamma_0 \end{aligned} \tag{18-124}$$

则式(18－123)变成

$$v_z(t) = v_z(t_0)\cos\Psi T + v_h(t_0)\sin\Psi T \tag{18-125}$$

定义

$$c_i \triangleq \cos(\xi_i T) \quad i = 1,2 \tag{18-126}$$

$$s_i \triangleq \sin(\xi_i T) \quad i = 1,2 \tag{18-127}$$

可以将式(18－119)和式(18－122)重写为

$$\begin{aligned} v_x(t) = \frac{1}{2}\{ & (c_1 + c_2)v_x(t_0) + (s_1 + s_2)v_y(t_0) + \\ & [(s_1 - s_2)\sin\vartheta_0 - (c_1 - c_2)\cos\vartheta_0]v_z(t_0)\} \end{aligned} \tag{18-128}$$

$$\begin{aligned} v_y(t) = \frac{1}{2}\{ & -(s_1 + s_2)v_x(t_0) + (c_1 + c_2)v_y(t_0) + \\ & [(c_1 - c_2)\sin\vartheta_0 + (s_1 - s_2)\cos\vartheta_0]v_z(t_0)\} \end{aligned} \tag{18-129}$$

但是，必须考虑有些特殊情况：

(1) 当 $\Omega = \Psi = 0$ 的情况。此时 $c_i = 1$ 且 $s_i = 0 (i = 1,2)$，得 $\boldsymbol{v}(t) = \boldsymbol{v}(t_0)$。

(2) 当 $\Omega = 0, \Psi \neq 0, \xi_1 = -\Psi$ 且 $\xi_2 = \Psi$ 时，得 $c_1 = c_2 = \cos(\Psi T)$，且 $s_1 = -\sin(\Psi T), s_2 = \sin(\Psi T)$。

(3) 当 $\Psi = 0, \Omega \neq 0$ 且 $\xi_1 = \xi_2 = \Omega$ 时，得 $c_1 = c_2 = \cos(\Omega T)$，且 $s_1 = s_2 = \sin(\Omega T)$。

(4) 如果 $\Omega = \Psi \neq 0$，有 $\xi_1 = 0, c_1 = 1$ 和 $s_1 = 0$。

(5) 如果 $\Omega = -\Psi \neq 0$，有 $\xi_2 = 0, c_2 = 1$ 和 $s_2 = 0$。

18. A. 2 恒定角速率偏转运动位置分量

为生成位置分量表达式 $\boldsymbol{r}(t)$，已知

$$\boldsymbol{r}(t) = \boldsymbol{r}(t_0) + \int_{t_0}^{t} \boldsymbol{v}(u)\mathrm{d}u \tag{18-130}$$

考虑积分的形式，即

$$\begin{aligned}
\int_{t_0}^{t} c_i \mathrm{d}u &= \int_{t_0}^{t} \cos(\xi_i(u - t_0))\mathrm{d}u \\
&= \int_{t_0}^{t} [\cos(\xi_i u)\cos(-\xi_i t_0) - \sin(\xi_i u)\sin(-\xi_i t_0)]\mathrm{d}u \\
&= \cos(-\xi_i t_0)\int_{t_0}^{t} \cos(\xi_i u)\mathrm{d}u - \sin(-\xi_i t_0)\int_{t_0}^{t} \sin(\xi_i u)\mathrm{d}u \\
&= \frac{1}{\xi_i}\sin(\xi_i t) \\
&= \frac{1}{\xi_i}s_i
\end{aligned} \tag{18-131}$$

类似地，有

$$\begin{aligned}
\int_{t_0}^{t} s_i \mathrm{d}u &= \int_{t_0}^{t} \sin(\xi_i(u - t_0))\mathrm{d}u \\
&= \frac{1}{\xi_i}(1 - \cos(\xi_i t)) \\
&= \frac{1}{\xi_i}(1 - c_i)
\end{aligned} \tag{18-132}$$

定义

$$g_i \triangleq \frac{1}{\xi_i}s_i \tag{18-133}$$

$$h_i \triangleq \frac{1}{\xi_i}(1 - c_i) \tag{18-134}$$

在执行积分之后，式(18－130)的分量由下式给出，即

$$\begin{aligned}
r_x(t) = r_x(t_0) + \frac{1}{2}\{&(g_1 + g_2)v_x(t_0) + (h_1 + h_2)v_y(t_0) + \\
&[(h_1 - h_2)\sin\vartheta_0 - (g_1 - g_2)\cos\vartheta_0]v_z(t_0)\}
\end{aligned} \tag{18-135}$$

$$\begin{aligned}
r_y(t) = r_y(t_0) + \frac{1}{2}\{&-(h_1 + h_2)v_x(t_0) + (g_1 + g_2)v_y(t_0) + \\
&[(g_1 - g_2)\sin\vartheta_0 + (h_1 - h_2)\cos\vartheta_0]v_z(t_0)\}
\end{aligned} \tag{18-136}$$

$$r_z(t) = r_z(t_0) + \frac{1 - \cos\Psi T}{\Psi}v_h(t_0) + \frac{\sin\Psi T}{\Psi}v_z(t_0) \tag{18-137}$$

考虑以下特殊情况：

(1) 考虑$\Omega=\Psi=0$情况。由于$\boldsymbol{v}(t)=\boldsymbol{v}(t_0)$,所以有$\boldsymbol{r}(t)=\boldsymbol{r}(t_0)+T\boldsymbol{v}(t_0)$。

(2) 当$\Omega=0,\Psi\neq0$时,得$g_1=g_2=\sin(\Psi T)/\Psi$和$h_1=-(1-\cos(\Psi T))/\Psi,h_2=-h_1$。

(3) 当$\Psi=0,\Omega\neq0$时,得$g_1=g_2=\sin(\Omega T)/\Omega$和$h_1=h_2=-(1-\cos(\Omega T))/\Omega$。

(4) 如果$\Omega=\Psi\neq0$,有$\xi_1=0$、$g_1=T$和$h_1=0$。

(5) 如果$\Omega=-\Psi\neq0$,有$\xi_2=0$、$g_2=1$和$h_2=0$。

18. A. 3 联合轨迹转移方程

将式(18-135)~式(18-137)、式(18-128)、式(18-129)和式(18-125)合并为一个矢量矩阵方程,得

$$\boldsymbol{x}(t_n)=\boldsymbol{f}_n(\boldsymbol{x}(t_0)) \tag{18-138}$$

$$\boldsymbol{x}(t_n)\triangleq\boldsymbol{x}_n=[r_{x,n},r_{y,n},r_{z,n},v_{x,n},v_{y,n},v_{z,n}]^{\mathrm{T}} \tag{18-139}$$

通过设置一个航途基准点,将航迹划分成段,并指定每段的转弯和俯冲速率,可用式(18-138)来生成真实航迹。对于这一案例中用到的所有仿真数据,生成航迹的时间间隔是0.1s或者采样频率为10Hz。

18. A. 4 基于期望偏转加速度来设置偏转速度

对于速度v恒定的目标,其速度分量是飞行方向和姿势的函数,由式(18-113)给出,加速度可以表示成

$$\boldsymbol{a}(t)=\frac{\mathrm{d}\boldsymbol{v}}{\mathrm{d}t}=\begin{bmatrix}a_x(t)\\a_y(t)\\a_z(t)\end{bmatrix}$$

$$=v\begin{bmatrix}\dot{\vartheta}\cos\vartheta(t)\cos\gamma(t)-\dot{\gamma}\sin\vartheta(t)\sin\gamma(t)\\-\dot{\vartheta}\sin\vartheta\cos\gamma(t)-\dot{\gamma}\cos\vartheta(t)\sin\gamma(t)\\\dot{\gamma}\cos\gamma(t)\end{bmatrix} \tag{18-140}$$

加速度的大小由下式给出,即

$$|\boldsymbol{a}(t)|=\sqrt{a_x^2(t)+a_y^2(t)+a_z^2(t)} \tag{18-141}$$

$$a_x^2(t)=v^2[\dot{\vartheta}^2\cos^2\vartheta(t)\cos^2\gamma(t)+\dot{\gamma}^2\sin^2\vartheta(t)\sin^2\gamma(t)-2\dot{\vartheta}\dot{\gamma}\cos\vartheta(t)\sin\vartheta(t)\cos\gamma(t)\sin\gamma(t)] \tag{18-142}$$

$$a_y^2(t)=v^2[\dot{\vartheta}^2\sin^2\vartheta(t)\cos^2\gamma(t)+\dot{\gamma}^2\cos^2\vartheta(t)\sin^2\gamma(t)+2\dot{\vartheta}\dot{\gamma}\cos\vartheta(t)\sin\vartheta(t)\cos\gamma(t)\sin\gamma(t)] \tag{18-143}$$

$$a_z^2(t)=v^2\dot{\gamma}^2\cos^2\gamma(t) \tag{18-144}$$

因此,有

$$|\boldsymbol{a}(t)| = v\sqrt{\dot{\vartheta}^2\cos\gamma(t) + \dot{\gamma}^2} \tag{18-145}$$

替换 $\dot{\vartheta} = \Omega$ 和 $\dot{\gamma} = \Psi$,上式可变为

$$|\boldsymbol{a}(t)| = v\sqrt{\Omega^2\cos\gamma(t) + \Psi^2} \tag{18-146}$$

考虑以下恒速率转弯/俯冲特殊情况:

(1) 对于水平转弯,$\gamma(t) \to \gamma_0$;$\Psi \to 0$ 且

$$|\boldsymbol{a}(t)| = \Omega v\cos\gamma_0 \tag{18-147}$$

(2) 对于垂直俯冲:$\Omega \to 0$ 且

$$|\boldsymbol{a}(t)| = \Psi v \tag{18-148}$$

假设指定一个水平转弯,其加速度定义为重力加速度 g 的倍数,如 $x\,g$。对于一个初始姿势 $\gamma_0 = 0$ 的水平转弯,有

$$\Omega = \frac{x}{v} \tag{18-149}$$

例如,指定一个以 600kn 的速度匀速运动的目标,以 $6g$ 加速度水平向右转弯,转弯速率应该设为

$$\Omega = \frac{6g}{600\text{kn}} \times \boldsymbol{C} \tag{18-150}$$

这里转换因子 C 定义为

$$C \triangleq \frac{[32.174(\text{ft/s}^2)/g]}{(6076.11549\text{n mile/ft})/3600(\text{s/h}))} \tag{18-151}$$

因此,对于这种情况,必须设置 $\Omega = 0.19\text{rad/s}$,以得所需要的转弯加速度。

附录 18.B 三维坐标变换

当使用一个跟踪滤波器,定义笛卡儿坐标系下状态矢量和球极坐标系下观测矢量,需要笛卡儿坐标系 —— 球坐标系和球坐标系 —— 笛卡儿坐标系的变换。这些转换必须应用到变换后的矢量以及协方差矩阵当中。本节将导出这些变换,并给出如何开发子程序以实现这些变换。作为附加好处,可以生成变换的雅可比矩阵,并成为输出的一部分。

给出多维坐标变换的一般概念。由一个坐标系统变换至另一个系统的矢量变换可以写为

$$\boldsymbol{r} = \boldsymbol{f}(\boldsymbol{x}) \tag{18-152}$$

式中:$\boldsymbol{x}$ 为原始坐标系中的矢量;$\boldsymbol{r}$ 为变换后坐标系中的矢量。

为给出协方差矩阵的变换,检验协方差矩阵在变换后系统中的定义,有

$$P^{rr} = \varepsilon\{(\boldsymbol{r}-\hat{\boldsymbol{r}})(\boldsymbol{r}-\hat{\boldsymbol{r}})^{\mathrm{T}}\} \tag{18-153}$$

这里 ε 代表期望值，$\hat{\boldsymbol{r}} = \varepsilon(\boldsymbol{r})$，由式(18 - 152)得

$$\hat{\boldsymbol{r}} = \boldsymbol{f}(\hat{\boldsymbol{x}}) \tag{18-154}$$

将 $\boldsymbol{f}(\boldsymbol{x})$ 展开为关于 $\hat{\boldsymbol{x}}$ 的一阶泰勒级数，由式(2 - 70)得

$$\begin{aligned}\boldsymbol{f}(\boldsymbol{x}) &= [\boldsymbol{f}(\boldsymbol{x})]_{x=\hat{x}} + [\boldsymbol{\nabla}_x \boldsymbol{f}^{\mathrm{T}}(\boldsymbol{x})]_{x=\hat{x}}^{\mathrm{T}}(\boldsymbol{x}-\hat{\boldsymbol{x}}) \\ &= \boldsymbol{f}(\hat{\boldsymbol{x}}) + \hat{\boldsymbol{F}}(\boldsymbol{x}-\hat{\boldsymbol{x}})\end{aligned} \tag{18-155}$$

这里将雅可比矩阵 $\hat{\boldsymbol{F}}$ 定义为

$$\hat{\boldsymbol{F}} = [\boldsymbol{\nabla}_x \boldsymbol{f}^{\mathrm{T}}(\boldsymbol{x})]_{x=\hat{x}}^{\mathrm{T}} = \begin{bmatrix} \dfrac{\partial f_1}{\partial x_1} & \dfrac{\partial f_1}{\partial x_2} & \cdots & \dfrac{\partial f_1}{\partial x_n} \\ \dfrac{\partial f_2}{\partial x_1} & \dfrac{\partial f_2}{\partial x_2} & \cdots & \dfrac{\partial f_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial f_m}{\partial x_1} & \dfrac{\partial f_m}{\partial x_2} & \cdots & \dfrac{\partial f_m}{\partial x_n} \end{bmatrix}_{x=\hat{x}} \tag{18-156}$$

式中：n 为矢量 $\boldsymbol{x}$ 的维数；m 为矢量 $\boldsymbol{f}$ 的维数。

将式(18 - 155)代入式(18 - 153)，得

$$\begin{aligned}\boldsymbol{P}^{rr} &= \varepsilon\{[\boldsymbol{f}(\boldsymbol{x}) - \boldsymbol{f}(\hat{\boldsymbol{x}})][\boldsymbol{f}(\boldsymbol{x}) - \boldsymbol{f}(\hat{\boldsymbol{x}})]^{\mathrm{T}}\} \\ &= \varepsilon\{[\boldsymbol{f}(\hat{\boldsymbol{x}}) + \hat{\boldsymbol{F}}(\boldsymbol{x}-\hat{\boldsymbol{x}}) - \boldsymbol{f}(\hat{\boldsymbol{x}})][\boldsymbol{f}(\hat{\boldsymbol{x}}) + \hat{\boldsymbol{F}}(\boldsymbol{x}-\hat{\boldsymbol{x}}) - \boldsymbol{f}(\hat{\boldsymbol{x}})]^{\mathrm{T}}\} \\ &= \hat{\boldsymbol{F}}\varepsilon\{(\boldsymbol{x}-\hat{\boldsymbol{x}})(\boldsymbol{x}-\hat{\boldsymbol{x}})^{\mathrm{T}}\}\boldsymbol{F}^{\mathrm{T}} \\ &= \hat{\boldsymbol{F}}\boldsymbol{P}^{xx}\boldsymbol{F}^{\mathrm{T}}\end{aligned} \tag{18-157}$$

式(18 - 157)给出了一阶协方差矩阵坐标变换的一般方程。注意，式(18 - 157)包括雅可比矩阵的计算，坐标变换子程序同时给出雅可比矩阵。

18. B. 1 笛卡儿坐标系变换为球坐标系

如图 18 - 25 所示，对于一般的三维坐标系统，可以将状态矢量写为笛卡儿坐标系形式，即

$$\boldsymbol{x} = [r_x, v_x, r_y, v_y, r_x, v_z]^{\mathrm{T}} \tag{18-158}$$

或者写为球极坐标系形式，即

$$\boldsymbol{r} = [R, \dot{R}, \theta, \dot{\theta}, \phi, \dot{\phi}]^{\mathrm{T}} \tag{18-159}$$

由图 18 - 25，可知从笛卡儿坐标系 —— 球坐标系的位置坐标变换为

$$\begin{bmatrix} R \\ \theta \\ \phi \end{bmatrix} = \begin{bmatrix} \sqrt{r_x^2 + r_y^2 + r_z^2} \\ \tan\left(\frac{r_x}{r_y}\right) \\ \tan\left(\frac{r_z}{\sqrt{r_x^2 + r_y^2}}\right) \end{bmatrix} \quad (18-160)$$

取每个分量关于时间的导数,得球速率分量为

$$\begin{bmatrix} \dot{R} \\ \dot{\theta} \\ \dot{\phi} \end{bmatrix} = \begin{bmatrix} \frac{r_x v_x + r_y v_y + r_z v_z}{R} \\ \frac{r_y v_x - r_x v_y}{R_h^2} \\ \frac{R v_z - r_z \dot{R}}{R R_h} \end{bmatrix} \quad (18-161)$$

这里水平距离定义为 $R_h \triangleq \sqrt{r_x^2 + r_y^2}$。在三维坐标系中,$R$ 通常是指斜线距离。这样,完整的笛卡儿坐标系——球坐标系的矢量变换 $\boldsymbol{h}(\boldsymbol{x})$ 定义为

$$\boldsymbol{r} = \begin{bmatrix} R \\ \dot{R} \\ \theta \\ \dot{\theta} \\ \phi \\ \dot{\phi} \end{bmatrix} = \boldsymbol{h}(\boldsymbol{x}) \triangleq \begin{bmatrix} h_1(\boldsymbol{x}) \\ h_2(\boldsymbol{x}) \\ h_3(\boldsymbol{x}) \\ h_4(\boldsymbol{x}) \\ h_5(\boldsymbol{x}) \\ h_6(\boldsymbol{x}) \end{bmatrix} = \begin{bmatrix} \sqrt{r_x^2 + r_y^2 + r_z^2} \\ \frac{r_x v_x + r_y v_y + r_z v_z}{R} \\ \arctan\left(\frac{r_x}{r_y}\right) \\ \frac{r_y v_x - r_x v_y}{R_H^2} \\ \arctan\left(\frac{r_z}{\sqrt{r_x^2 + r_y^2}}\right) \\ \frac{R v_z - r_z \dot{R}}{R R_h} \end{bmatrix} \quad (18-162)$$

由式(18 - 156),可以将笛卡儿坐标系——球坐标系协方差变换的雅可比矩阵写为

$$\hat{\boldsymbol{H}} = \begin{bmatrix} H_{11} & H_{12} & \cdots & H_{16} \\ H_{21} & H_{22} & \cdots & H_{26} \\ \vdots & \vdots & \ddots & \vdots \\ H_{61} & H_{62} & \cdots & H_{66} \end{bmatrix} = \begin{bmatrix} \frac{\partial h_1}{\partial r_x} & \frac{\partial h_1}{\partial v_x} & \cdots & \frac{\partial h_1}{\partial v_z} \\ \frac{\partial h_2}{\partial r_x} & \frac{\partial h_1}{\partial v_x} & \cdots & \frac{\partial h_2}{\partial v_z} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial h_6}{\partial r_x} & \frac{\partial h_6}{\partial v_x} & \cdots & \frac{\partial h_6}{\partial v_z} \end{bmatrix}_{\boldsymbol{x}=\hat{\boldsymbol{x}}} \quad (18-163)$$

如果先计算式(18 - 162),在每步计算中可以使用球坐标分量,得雅可比分量为

$$H_{11} = \frac{\partial}{\partial r_x}[(r_x^2 + r_y^2 + r_z^2)^{1/2}] = \frac{r_x}{R} \tag{18-164}$$

$$H_{12} = H_{14} = H_{16} = 0 \tag{18-165}$$

$$H_{13} = \frac{r_y}{R} \tag{18-166}$$

$$H_{15} = \frac{r_z}{R} \tag{18-167}$$

$$H_{21} = \frac{Rv_x - r_x\dot{R}}{R^2} \tag{18-168}$$

$$H_{22} = H_{11} \tag{18-169}$$

$$H_{23} = \frac{Rv_y - r_y\dot{R}}{R^2} \tag{18-170}$$

$$H_{24} = H_{13} \tag{18-171}$$

$$H_{25} = \frac{Rv_z - r_z\dot{R}}{R^2} \tag{18-172}$$

$$H_{26} = H_{15} \tag{18-173}$$

$$H_{31} = \frac{r_y}{R_h^2} \tag{18-174}$$

$$H_{32} = H_{34} = H_{35} = H_{36} = 0 \tag{18-175}$$

$$H_{33} = -\frac{r_x}{R_h^2} \tag{18-176}$$

$$H_{41} = -\frac{v_y - 2r_x\dot{\theta}}{R_h^2} \tag{18-177}$$

$$H_{42} = H_{31} \tag{18-178}$$

$$H_{43} = \frac{v_x - 2r_y\dot{\theta}}{R_h^2} \tag{18-179}$$

$$H_{44} = H_{33} \tag{18-180}$$

$$H_{45} = H_{46} = 0 \tag{18-181}$$

$$H_{51} = -\frac{r_x r_z}{R^2 R_h} \tag{18-182}$$

$$H_{52} = H_{54} = H_{56} = 0 \tag{18-183}$$

$$H_{53} = -\frac{r_y r_z}{R^2 R_h} \tag{18-184}$$

$$H_{55} = \frac{R_h}{R^2} \tag{18-185}$$

$$H_{61} = \frac{1}{RR_h}\left[\frac{r_x}{R_h^2}(r_z\dot{R} - Rv_z) + \frac{r_z}{R^2}(2r_x\dot{R} - Rv_x)\right] \tag{18 - 186}$$

$$H_{62} = H_{51} \tag{18 - 187}$$

$$H_{63} = \frac{1}{RR_h}\left[\frac{r_y}{R_h^2}(r_z\dot{R} - Rv_z) + \frac{r_z}{R^2}(2r_y\dot{R} - Rv_y)\right]$$

$$H_{64} = H_{53} \tag{18 - 188}$$

$$H_{65} = \frac{1}{R^3R_h}[(2r_z^2 + R^2)\dot{R} - r_zRv_z] \tag{18 - 189}$$

$$H_{66} = H_{55} \tag{18 - 190}$$

利用式(18 - 157),可以将笛卡儿坐标系 —— 球坐标系的协方差矩阵变换写成

$$\boldsymbol{P}^{rr} = \hat{\boldsymbol{H}}\boldsymbol{P}^{xx}\hat{\boldsymbol{H}}^{\mathrm{T}} \tag{18 - 191}$$

式中:$\boldsymbol{P}^{xx}$ 和 $\boldsymbol{P}^{rr}$ 分别为笛卡儿坐标系和球坐标系下的协方差矩阵。

18. B. 2　球坐标系变换为笛卡儿坐标系

有时需要进行式(18 - 162) 的逆运算,即球坐标系 —— 笛卡儿坐标系变换的定义为

$$\boldsymbol{x} = \boldsymbol{g}(\boldsymbol{r}) \tag{18 - 192}$$

首先检验位置分量,使用图 18 - 25,得

$$\begin{bmatrix} r_x \\ r_y \\ r_z \end{bmatrix} = \begin{bmatrix} R\sin\theta\cos\phi \\ R\cos\theta\cos\phi \\ R\sin\phi \end{bmatrix} \tag{18 - 193}$$

取每个分量的导数,得

$$\begin{bmatrix} v_x \\ v_y \\ v_z \end{bmatrix} = \begin{bmatrix} \sin\theta\cos\phi & R\cos\theta\cos\phi & -R\sin\theta\sin\phi \\ \cos\theta\cos\phi & -R\sin\theta\cos\phi & -R\cos\theta\sin\phi \\ \sin\phi & 0 & R\cos\phi \end{bmatrix}\begin{bmatrix} \dot{R} \\ \dot{\theta} \\ \dot{\phi} \end{bmatrix} \tag{18 - 194}$$

合并式(18 - 193) 和式(18 - 194),得球坐标系 —— 笛卡儿坐标系的变换为

$$\boldsymbol{x} = \begin{bmatrix} r_x \\ v_x \\ r_y \\ v_y \\ r_z \\ v_z \end{bmatrix} = \begin{bmatrix} g_1(\boldsymbol{r}) \\ g_2(\boldsymbol{r}) \\ g_3(\boldsymbol{r}) \\ g_4(\boldsymbol{r}) \\ g_5(\boldsymbol{r}) \\ g_6(\boldsymbol{r}) \end{bmatrix} = \begin{bmatrix} R\sin\theta\cos\phi \\ \dot{R}\sin\theta\cos\phi + \dot{\theta}R\cos\theta\cos\phi - \dot{\phi}R\sin\theta\sin\phi \\ R\cos\theta\cos\phi \\ \dot{R}\cos\theta\cos\phi - \dot{\theta}R\sin\theta\cos\phi - \dot{\phi}R\cos\theta\sin\phi \\ R\sin\phi \\ \dot{R}\sin\phi + \dot{\phi}R\cos\phi \end{bmatrix} \tag{18 - 195}$$

给出协方差矩阵由球坐标系 —— 笛卡儿坐标系变换方程为

$$\boldsymbol{P}^{xx}=\hat{\boldsymbol{G}}\boldsymbol{P}^{rr}\hat{\boldsymbol{G}}^{\mathrm{T}} \tag{18-196}$$

式(18－156)的雅可比变换分量为

$$G_{11}=\sin\theta\cos\phi \tag{18-197}$$

$$G_{12}=G_{14}=G_{16}=0 \tag{18-198}$$

$$G_{13}=R\cos\theta\cos\phi \tag{18-199}$$

$$G_{15}=-R\sin\theta\sin\phi \tag{18-200}$$

$$G_{21}=\dot{\theta}\cos\theta\cos\phi-\dot{\phi}\sin\theta\sin\phi \tag{18-201}$$

$$G_{22}=G_{11} \tag{18-202}$$

$$G_{23}=\dot{R}\cos\theta\cos\phi-\dot{\theta}R\sin\theta\cos\phi-\dot{\phi}R\cos\theta\sin\phi \tag{18-203}$$

$$G_{24}=G_{13} \tag{18-204}$$

$$G_{25}=-\dot{R}\sin\theta\sin\phi-\dot{\theta}R\cos\theta\sin\phi-\dot{\phi}R\sin\theta\cos\phi \tag{18-205}$$

$$G_{26}=G_{15} \tag{18-206}$$

$$G_{31}=\cos\theta\cos\phi \tag{18-207}$$

$$G_{32}=G_{34}=G_{36}=0 \tag{18-208}$$

$$G_{33}=-R\sin\theta\cos\phi \tag{18-209}$$

$$G_{35}=-R\cos\theta\sin\phi \tag{18-210}$$

$$G_{41}=-\dot{\theta}\sin\theta\cos\phi-\dot{\phi}\cos\theta\sin\phi \tag{18-211}$$

$$G_{42}=G_{31} \tag{18-212}$$

$$G_{43}=-\dot{R}\sin\theta\cos\phi-\dot{\theta}R\cos\theta\cos\phi+\dot{\phi}R\sin\theta\sin\phi \tag{18-213}$$

$$G_{44}=G_{33} \tag{18-214}$$

$$G_{45}=-\dot{R}\cos\theta\sin\phi+\dot{\theta}R\sin\theta\sin\phi-\dot{\phi}R\cos\theta\cos\phi \tag{18-215}$$

$$G_{46}=G_{35} \tag{18-216}$$

$$G_{51}=\sin\phi \tag{18-217}$$

$$G_{52}=G_{53}=G_{54}=G_{56}=0 \tag{18-218}$$

$$G_{55}=R\cos\phi \tag{18-219}$$

$$G_{61}=\dot{\phi}\cos\phi \tag{18-220}$$

$$G_{62}=G_{51} \tag{18-221}$$

$$G_{63}=G_{64}=0 \tag{18-222}$$

$$G_{65}=\dot{R}\cos\phi-\dot{\phi}R\sin\phi \tag{18-223}$$

$$G_{66}=G_{55} \tag{18-224}$$

参考文献

[1] Blackman SS. Multiple - Target Tracking with Radar Applications. Artech House, 1986.

[2] Bar - Shalom Y, Fortman TE. Tracking and Data Association. Academic Press, 1988.

[3] Bar - Shalom Y, Rong Li X, Kirubarajan T. Estimation with Application to Tracking and Navigation. Wiley, 2001.

[4] Ristic B, Arulampalam S, Gordon N. Beyond the Kalmanfilter: Particle Filter Applications. Artech House, 2004.

[5] Zollo S, Ristic B. On the Choice of the Coordinate System and Tracking Filter for the Track - while - scan Mode of an Airborne Radar. DSTO - TR - 0926, 1999.

[6] Lefferts RE. Alpha - beta filters in polar coordinates with acceleration constraints. IEEE Trans. Aero. Elec. Syst, 1988, 24(6): 693 - 699.

[7] Blair WD, Watson GA, Rice TR. Interacting multiple model filter for tracking maneuvering targets in spherical coordinates. Proc. Am. Control Conf, 1991: 1055 - 1059.

[8] Kameda H, Nomoto K, Kosuge Y, Kondo M. Target tracking algorithm in radar reference coordinates using fully - coupled filters. Electron. Commun. Jpn, 1997, 10(7): 327 - 333.

[9] Swillinger DZ. CRC Standard Mathematical Tables and Formulae. 30th ed. CRC Press, 1996.

第19章 利用摄影测量法跟踪下落刚体

19.1 概述

摄影测量法属于数学科学范畴,即:用多个摄像机获取二维视频图像帧数据序列,从中观测得一组目标特性,在此基础上估计目标的姿势(三维位置和方向)和运动属性(平移、旋转的速度和加速度)。人们提出许多方法来求解这个估计问题,1990 年之前最常用的是非线性最小二乘法(nonlinear least squares,NLLSQ)方法。Huang 和 Netravali 在文献[1]中对早期的 NLLSQ 方法进行了综述,在 NLLSQ 方法的每个时间步中,用预测视频图像与当前实际视频图像之间的残差来进行一个迭代的 NLLSQ 过程,使得局部残余误差最小化,进而得到刚体状态估计。另一种更为现代的算法是贝叶斯估计扩展卡尔曼滤波器(EKF)方法,由 Iu 和 Wohn[2] 和 Gennery[3] 提出,并得了非常好的试验结果。其后,Welch[4,5] 提出一种 SCAAT EKF 方法。

上面提到的方法都有一些缺点,会影响刚体状态估计性能。NLLSQ 方法假定动态模型中没有噪声,在时间上不递归,这样每一次迭代都需要计算雅可比矩阵(偏导数矩阵)。观测值集合由视频帧图像组成,每次迭代都需要重新处理的所有图像帧,因此该方法非常耗时。而且,因为应用于摄影测量中的 NLLSQ 是一种"一次一步"方法,在 NLLSQ 中不估计平移速度和角度变化率,必须使用后处理平滑步骤单独计算。

为了使观测过程线性化(见第7章),每个时间步必须计算非线性观测模型和动态转移模型的雅可比矩阵,因此 EKF 方法很难实现。已经证明,对于很多高度非线性问题,这种解析线性化有很多不足。

为进行安全性评估,本章致力于研究估计和跟踪中的特殊问题——估计和跟踪从飞机外挂架上落下的刚体(看似炸弹的目标)随时间变化的姿势。这里所

使用的图像序列是由多个附属于同一参考目标的高速图像传感器记录得。用图像传感器或摄像机可靠观测跟踪目标的刚体状态轨迹。图19－1和图19－2给出标记在参考目标和刚体目标上的特征点。在飞机上共装有19个摄像机，其中翼尖上面有4个摄像机，飞机尾部下面4个摄像机，机头下面1个摄像机。每个摄像机的图像帧中，特征点的位置都记录在一个时域图像序列当中，用于估计跟踪目标的航迹。每个图像传感器由特定视角记录目标运动，得到各自独特的图像序列。每个图像序列必须适当地结合其它摄像机的图像序列，在每个时间步计算航迹的最佳跟踪估计。跟踪目标的航迹可以用来进行以安全为目的的脱靶量分析，或对其他类型的建模与仿真进行验证和改进。因为图像传感器估计3D空间中刚体目标随时间变化的位置和方向，将摄影测量法估计用于准确近距离跟踪是首选方案[6]。



图19－1　美国海军F－18E超级大黄蜂式飞机准备投放4个Mk－62水雷



图19－2　由安装在飞机尾部下面的摄像机拍摄到的投放MK－62的特写镜头

在估计过程中，本书提出两种新方法，即无迹卡尔曼滤波器（UKF）和广义UKF粒子滤波器（分别参见第9章和第17章）。因为这些滤波器能够估计刚体运动状态，包括笛卡儿速度和加速度以及角速度和角加速度，而无需计算雅可比矩阵，因此性能优于EKF和NLLSQ。

第19.2节提出一种动态运动模型，以描述下落刚体目标的平移和旋转运

动。该节还讨论了动态噪声项,用来描述动态模型中微小的运动偏离。第 19.3 节介绍了观测模型,首先计算刚体表面特征点的位置,将刚体上的三维特征点由本体坐标系变换(旋转和平移)至参考(飞机的)坐标系;然后将三维点变换至摄像机的坐标系并将三维点投影到摄像机的二维图像平面。

第 19.4 节讨论了所使用的估计方法,第 19.4.1 节为 NLLSQ,第 19.4.2 节为 UKF,19.4.3 为 UKF 组合粒子滤波器,第 19.4.5 节给出所有滤波器的初始值。

为进行性能分析,给出一种综合刚体模型和综合运动模型,以模拟刚体由飞机外挂架投放后的运动。第 19.5 节用这个模型和方法来生成模拟图像帧数据。第 19.6 节给出所有估计方法相关性能,并进行了重要结果总结。

附录 19.A 概括介绍了一些定义和相关数学概念,包括四元数、轴角矢量和旋转等。

19.2 刚体运动的过程模型

这里运动模型为时间函数,描述刚体的平移和旋转运动。假定一个瞬时状态矢量 $\boldsymbol{x}(t) = [\boldsymbol{p}^{\mathrm{T}}(t), \boldsymbol{a}^{\mathrm{T}}(t)]^{\mathrm{T}}$ 定义了相对于固定参考笛卡儿坐标系中图像帧的刚体重心的平移 $\boldsymbol{p}(t) = [x(t), y(t), z(t)]^{\mathrm{T}}$,并用一个轴角矢量 $\boldsymbol{a}(t) = [a_x(t), a_y(t), a_z(t)]^{\mathrm{T}}$ 来表示刚体相对于参考帧的方向。在某些条件下,矢量 $\boldsymbol{a}$ 可能变得数值上不稳定(尤其是当 $\|\boldsymbol{a}\| = 2\pi n$ 时,其中 n 为正整数),所以用单位四元数 $\boldsymbol{q} = [q_s, q_x, q_y, q_z]^{\mathrm{T}}$ 与 $\boldsymbol{a}$ 结合来代表方向。在第 19.2.2 节和附录 19.A.2 中,详细讨论了 $\boldsymbol{q}$ 值的计算。在后面给出的状态模型(过程)中,将调整状态矢量,用速度矢量、加速度矢量和加加速度(三次时间导数)项来控制平移和旋转运动。

因为刚体的平移运动与旋转运动无关,将分别处理这两种运动。在下面章节将给出动态转移方程,将上面定义的刚体状态矢量从时刻 t_{n-1} 转移至之后的 t_n。在第 19.2.1 节,用转移方程控制状态矢量的平移部分,在第 19.2.2 节,用转移方程处理状态矢量的方向部分。一般来说,可以将总体转移方程写成如下形式,即

$$\boldsymbol{x}_n = \boldsymbol{f}(\boldsymbol{x}_{n-1}, \boldsymbol{v}_{n-1}) \tag{19-1}$$

式中:$\boldsymbol{v}_{n-1}$ 为驱动过程的噪声。在所有情况下都把噪声视为高斯加性噪声,这样式(19-1)变成

$$\boldsymbol{x}_n = \boldsymbol{f}(\boldsymbol{x}_{n-1}) + \boldsymbol{v}_{n-1} \tag{19-2}$$

第 19.2.3 节深入讨论了这种组合的姿势转移模型。第 19.2.4 节将讨论 $\boldsymbol{v}_{n-1}$ 的噪声模型的性质。

19.2.1 刚体平移运动的动态转移

令 $\boldsymbol{p}_n$ 为一个三维矢量,代表 t_n 时刻刚体重心(center of gravity,cg)的笛卡

儿位置坐标，即$\boldsymbol{p}_n = [x_n, y_n, z_n]^{\mathrm{T}}$。将$\boldsymbol{p}(t_n)$展开为关于$t_{n-1}$时刻的时域泰勒展开式，有

$$\boldsymbol{p}_n = \boldsymbol{p}_{n-1} + \left[\frac{\mathrm{d}\boldsymbol{p}(t)}{\mathrm{d}t}\right]_{t=t_{n-1}}(t_n - t_{n-1}) + \left[\frac{\mathrm{d}^2\boldsymbol{p}(t)}{\mathrm{d}t^2}\right]_{t=t_{n-1}}\frac{(t_n - t_{n-1})^2}{2!} + \left[\frac{\mathrm{d}^3\boldsymbol{p}(t)}{\mathrm{d}t^3}\right]_{t=t_{n-1}}\frac{(t_n - t_{n-1})^3}{3!} + \cdots \tag{19-3}$$

$$= \boldsymbol{p}_{n-1} + T_n\dot{\boldsymbol{p}}_{n-1} + \frac{T_n^2}{2}\ddot{\boldsymbol{p}}_{n-1} + \frac{T_n^3}{6}\dddot{\boldsymbol{p}}_{n-1} + \cdots \tag{19-4}$$

这里$\dot{\boldsymbol{p}}$、$\ddot{\boldsymbol{p}}$和$\dddot{\boldsymbol{p}}$代表前三阶位置的时间导数，$T_n \triangleq t_n - t_{n-1}$。这个例子中，假设所有观测摄像机同步记录图像，那么可假定T_n为常数T。仅保留至三阶项，将式(19-3)改写为

$$\boldsymbol{p}_n = \boldsymbol{p}_{n-1} + T\dot{\boldsymbol{p}}_{n-1} + \frac{T^2}{2}\ddot{\boldsymbol{p}}_{n-1} + \frac{T^3}{6}\dddot{\boldsymbol{p}}_{n-1} \tag{19-5}$$

对于恒定位置模型，$\dot{\boldsymbol{p}}$、$\ddot{\boldsymbol{p}}$和$\dddot{\boldsymbol{p}}$均为零。因此定义刚体的零阶三维状态矢量中仅包含位置分量$\boldsymbol{p}_n^{(0)} = \boldsymbol{p}_n = [x(t), y(t), z(t)]^{\mathrm{T}}$，式(19-5)变为

$$\boldsymbol{p}_n^{(0)} = \boldsymbol{p}_{n-1}^{(0)} + \boldsymbol{v}_{n-1}^{(0)} \tag{19-6}$$

这里增加了一个零阶过程速度噪声项$\boldsymbol{v}_{n-1}^{(0)}$。增加过程噪声是为了表示模型中由于一些未建模因素如风力、摩擦等导致的目标位置的微小偏差。

基于速度恒定假设，定义一个包含位置和速度分量的6维状态矢量$\boldsymbol{p}_n^{(1)} = [\boldsymbol{p}_n^{\mathrm{T}}, \dot{\boldsymbol{p}}_n^{\mathrm{T}}]^{\mathrm{T}}$，则式(19-5)变为一阶过程模型，即

$$\boldsymbol{p}_n^{(1)} = \boldsymbol{F}^{(1)}\boldsymbol{p}_{n-1}^{(1)} + \boldsymbol{v}_{n-1}^{(1)} \tag{19-7}$$

$$\boldsymbol{F}^{(1)} = \begin{bmatrix} \boldsymbol{I}_3 & T\boldsymbol{I}_3 \\ \boldsymbol{0}_3 & \boldsymbol{I}_3 \end{bmatrix} \tag{19-8}$$

式中：$\boldsymbol{I}_3$和$\boldsymbol{0}_3$分别为三维单位矩阵和零矩阵；$\boldsymbol{v}_{n-1}^{(1)}$为一阶加速度过程噪声。

采用相似的形式，对于加速度恒定模型(19-5)，9维状态矢量$\boldsymbol{p}_n^{(2)} = [\boldsymbol{p}_n^{\mathrm{T}}, \dot{\boldsymbol{p}}_n^{\mathrm{T}}, \ddot{\boldsymbol{p}}_n^{\mathrm{T}}]^{\mathrm{T}}$和过程模型变成

$$\boldsymbol{p}_n^{(2)} = \boldsymbol{F}^{(2)}\boldsymbol{p}_{n-1}^{(2)} + \boldsymbol{v}_{n-1}^{(2)} \tag{19-9}$$

$$\boldsymbol{F}^{(2)} = \begin{bmatrix} \boldsymbol{I}_3 & \boldsymbol{I}_3 & \frac{T^2}{2}\boldsymbol{I}_3 \\ \boldsymbol{0}_3 & \boldsymbol{I}_3 & T\boldsymbol{I}_3 \\ \boldsymbol{0}_3 & \boldsymbol{0}_3 & \boldsymbol{I}_3 \end{bmatrix} \tag{19-10}$$

式中：$\boldsymbol{v}_{n-1}^{(2)}$为二阶加加速度过程噪声。

对于恒定加加速度的12维模型，$\boldsymbol{p}_n^{(3)} = [\boldsymbol{p}_n^{\mathrm{T}}, \dot{\boldsymbol{p}}_n^{\mathrm{T}}, \ddot{\boldsymbol{p}}_n^{\mathrm{T}}, \dddot{\boldsymbol{p}}_n^{\mathrm{T}}]^{\mathrm{T}}$，进而得过程模

型为

$$\boldsymbol{p}_{n}^{(3)} = \boldsymbol{F}^{(3)}\boldsymbol{p}_{n-1}^{(3)} + \boldsymbol{v}_{n-1}^{(3)} \tag{19-11}$$

$$\boldsymbol{F}^{(3)} = \begin{bmatrix} \boldsymbol{I}_3 & T\boldsymbol{I}_3 & \frac{T^2}{2}\boldsymbol{I}_3 & \frac{T^3}{6}\boldsymbol{I}_3 \\ \boldsymbol{0}_3 & \boldsymbol{I}_3 & T\boldsymbol{I}_3 & \frac{T^2}{2}\boldsymbol{I}_3 \\ \boldsymbol{0}_3 & \boldsymbol{0}_3 & \boldsymbol{I}_3 & T\boldsymbol{I}_3 \\ \boldsymbol{0}_3 & \boldsymbol{0}_3 & \boldsymbol{0}_3 & \boldsymbol{I}_3 \end{bmatrix} \tag{19-12}$$

式中：$\boldsymbol{v}_{n-1}^{(1)}$ 为三阶 snap（或 jounce）四阶导数过程噪声。

19.2.2 刚体旋转运动的动态转移

可以用多种方式表示刚体的瞬时方向。这个方向可以表示为一组相对于 3 个笛卡儿参考帧轴的欧拉旋转角。这里选择的右手笛卡儿坐标系参考帧是 x 轴指向前方，y 轴当向前看时水平向右（右舷），z 轴竖直向下。3 个方向角为 $\{\phi,\theta,\psi\}$，其中 ϕ（翻滚）表示相对于 x 轴逆时针（counterclockwise，CCW）旋转，θ（俯仰）表示相对于 y 轴逆时针旋转，φ（偏转）表示相对于 z 轴逆时针旋转。但是，利用欧拉角描述刚体方向在某些情况下会导致奇异点[3]。

19.2.2.1 旋转表达式之间的转换

将 $\boldsymbol{a} = [a_x, a_y, a_z]^{\mathrm{T}} \in \mathbf{R}^3$ 定义为旋转的轴角表示，这里 $\boldsymbol{a}$ 的方向指定旋转轴的角度，$\|\boldsymbol{a}\| = \sqrt{a_x^2 + a_y^2 + a_z^2}$ 表示旋转角度（用弧度表示）。

令 $\boldsymbol{q} = [q_s, q_x, q_y, q_z]^{\mathrm{T}} \in \boldsymbol{H}$，其中 $\boldsymbol{H}$ 是四元数空间，定义为旋转的单位四元数表示。这里 q_s 为实数，q_x、q_y 和 q_z 为纯虚数系数，并且满足 $\sqrt{q_s^2 + q_x^2 + q_y^2 + q_z^2} = 1$。

令 $\boldsymbol{M} \in \mathbf{R}^3 \times \mathbf{R}^3$ 为一个 3×3 标准正交旋转矩阵。

为将轴角矢量 $\boldsymbol{a}$ 转换为单位四元数 $\boldsymbol{q}$，定义函数 $\mathcal{Q}_{\boldsymbol{a}}: \mathbf{R}^3 \to \boldsymbol{H}$，则有

$$\boldsymbol{q} = \mathcal{Q}_{\boldsymbol{a}}(\boldsymbol{a}) = \left[\cos\frac{\|\boldsymbol{a}\|}{2}, \frac{\sin\frac{\|\boldsymbol{a}\|}{2}}{\|\boldsymbol{a}\|}[a_x, a_y, a_z]\right]^{\mathrm{T}} \tag{19-13}$$

为将单位四元数 $\boldsymbol{q}$ 转换为角轴矢量 $\boldsymbol{a}$，定义函数 $\mathcal{A}_{\boldsymbol{q}}: \boldsymbol{H} \to \mathbf{R}^3$，则有

$$\boldsymbol{a} = \mathcal{A}_{\boldsymbol{q}}(\boldsymbol{q}) = \frac{2\cos^{-1} q_s}{\sqrt{1 - q_s^2}}[q_x, q_y, q_z]^{\mathrm{T}} \tag{19-14}$$

19.2.2.2 四元数乘法

设 $\boldsymbol{q}_1$、$\boldsymbol{q}_2$ 和 $\boldsymbol{q}_3$ 为四元数，四元数乘法 $\boldsymbol{q}_3 = \boldsymbol{q}_1\boldsymbol{q}_2$ 定义为

$$\boldsymbol{q}_3 = \begin{bmatrix} q_{3,s} \\ q_{3,x} \\ q_{3,y} \\ q_{3,z} \end{bmatrix} = \begin{bmatrix} q_{1,s} & -q_{1,x} & -q_{1,y} & -q_{1,z} \\ q_{1,x} & q_{1,s} & -q_{1,z} & q_{1,y} \\ q_{1,y} & q_{1,z} & q_{1,s} & -q_{1,x} \\ q_{1,z} & -q_{1,y} & q_{1,x} & q_{1,s} \end{bmatrix} \begin{bmatrix} q_{2,s} \\ q_{2,x} \\ q_{2,y} \\ q_{2,z} \end{bmatrix} \tag{19-15}$$

19.2.2.3 方向的动态模型

利用轴角表达式来表示刚体方向的动态时间转移。轴角方向矢量 $\boldsymbol{a}_n$ 表示相对于之前方向的一个微小旋转，即从 t_{n-1} 到 t_n 时刻的旋转。将这个微小轴角旋转加入（通过旋转）明确的外部四元数 $\boldsymbol{q}_{n-1}$，其中包含用在观测模型中的 t_{n-1} 时刻的完整旋转。注意，$\boldsymbol{a}$ 及其导数 $\dot{\boldsymbol{a}}$、$\ddot{\boldsymbol{a}}$ 和 $\dddot{\boldsymbol{a}}$（如果存在的话）都是相对参考坐标系帧而不是本体坐标系帧，因此在动态模型中不需要出现 $\boldsymbol{q}$ 值。

定义状态矢量方向上的轴角矢量 $\boldsymbol{a}$ 及其导数。恒定方向模型中 $\dot{\boldsymbol{a}}$、$\ddot{\boldsymbol{a}}$ 和 $\dddot{\boldsymbol{a}}$ 均为零，因此将零阶模型的状态矢量看成三维轴角矢量 $\boldsymbol{a}^{(0)}$，有

$$\boldsymbol{a}_n^{(0)} = \boldsymbol{a}_{n-1}^{(0)} + \boldsymbol{\rho}_{n-1}^{(0)} \tag{19-16}$$

式中：$\boldsymbol{\rho}_{n-1}^{(0)}$ 为方向的三维零阶角速度噪声。

对于恒定旋转速率模型，选择 6 维状态矢量 $\boldsymbol{a}_n^{(1)} \triangleq [\boldsymbol{a}_n^{\mathrm{T}}, \dot{\boldsymbol{a}}_n^{\mathrm{T}}]^{\mathrm{T}}$，得一阶过程模型为

$$\boldsymbol{a}_n^{(1)} = \boldsymbol{g}_{n-1}^{(1)}(\boldsymbol{a}_{n-1}^{(1)}) + \boldsymbol{\rho}_{n-1}^{(1)} \tag{19-17}$$

式中：$\rho_{n-1}^{(1)}$ 为方向的三维一阶角速度噪声。为将轴角矢量 $\boldsymbol{a}$ 在时间区间 $T = t_n - t_{n-1}$ 向前传播，不能简单地把 $T\dot{\boldsymbol{a}}_{n-1}$ 加入 $\boldsymbol{a}_{n-1}$ 来得一个新的矢量 $\boldsymbol{a}_n$，因为不能用矢量轴角表达式相加来得两个三维旋转的叠加（或组合）。反之，把轴角矢量变换为单位四元数，计算它们的乘积（记得颠倒它们的顺序），然后把结果变换回轴角矢量（或者，可以用旋转矩阵代替单位四元数）。具体而言，为 $\Delta\boldsymbol{a}$ 是方向的递增变化，由时间区间 T 内恒定旋转速率产生，因此有

$$\Delta\boldsymbol{a} = \int_{t_{n-1}}^{t_n} \dot{\boldsymbol{a}}\mathrm{d}t = T\dot{\boldsymbol{a}} \tag{19-18}$$

将 $\boldsymbol{q}_{\boldsymbol{a}_{n-1}}$ 和 $\boldsymbol{q}_{\Delta\boldsymbol{a}}$ 定义为

$$\boldsymbol{q}_{\boldsymbol{a}_{n-1}} = \mathcal{Q}_{\boldsymbol{a}}(\boldsymbol{a}_{n-1}) \tag{19-19}$$

$$\boldsymbol{q}_{\Delta\boldsymbol{a}} = \mathcal{Q}_{\boldsymbol{a}}(\Delta\boldsymbol{a}) = \mathcal{Q}_{\boldsymbol{a}}(T\dot{\boldsymbol{a}}) \tag{19-20}$$

将这两个四元数相乘，记得颠倒它们的顺序，将结果变换回轴角矢量 $\boldsymbol{a}_n$，即

$$\boldsymbol{a}_n = \mathcal{A}_q(\boldsymbol{q}_{\Delta\boldsymbol{a}}\boldsymbol{q}_{\boldsymbol{a}_{n-1}}) \tag{19-21}$$

假设转速恒定，所以有

$$\dot{\boldsymbol{a}}_n = \dot{\boldsymbol{a}}_{n-1} \tag{19-22}$$

现在，将式（19-19）~ 式（19-22）代入式（19-17），得

$$\boldsymbol{a}_n^{(1)} = \begin{bmatrix} \boldsymbol{a}_n \\ \dot{\boldsymbol{a}}_n \end{bmatrix} = \begin{bmatrix} \mathcal{A}_q(\mathcal{Q}_a(T\dot{\boldsymbol{a}}_{n-1})\mathcal{Q}_a(\boldsymbol{a}_{n-1})) \\ \dot{\boldsymbol{a}}_{n-1} \end{bmatrix} + \boldsymbol{\rho}_{n-1}^{(1)} \tag{19-23}$$

对于一个恒定旋转加速度模型，选择9维方向状态矢量 $\boldsymbol{a}_n^{(2)} \triangleq [\boldsymbol{a}_n^{\mathrm{T}}, \dot{\boldsymbol{a}}_n^{\mathrm{T}}, \ddot{\boldsymbol{a}}_n^{\mathrm{T}}]^{\mathrm{T}}$，得到二阶过程模型为

$$\boldsymbol{a}_n^{(2)} = \boldsymbol{g}_{n-1}^{(2)}(\boldsymbol{a}_{n-1}^{(2)}) + \boldsymbol{\rho}_{n-1}^{(2)} \tag{19-24}$$

式中：$\rho_{n-1}^{(2)}$ 为方向的三维二阶角速度噪声。在这个模型中，时间区间 T 方向的递增变化为

$$\Delta\boldsymbol{a} = \int_{t_{n-1}}^{t_n} (\dot{\boldsymbol{a}} + t\ddot{\boldsymbol{a}})\mathrm{d}t = T\dot{\boldsymbol{a}} + \frac{T^2}{2}\ddot{\boldsymbol{a}} \tag{19-25}$$

由于 $\dot{\boldsymbol{a}}$ 和 $\ddot{\boldsymbol{a}}$ 为特征矢量，所以有

$$\dot{\boldsymbol{a}}_n = \dot{\boldsymbol{a}}_{n-1} + T\ddot{\boldsymbol{a}}_{n-1} \tag{19-26}$$

$$\ddot{\boldsymbol{a}}_n = \ddot{\boldsymbol{a}}_{n-1} \tag{19-27}$$

因此有

$$\begin{bmatrix} \boldsymbol{a}_n \\ \dot{\boldsymbol{a}}_n \\ \ddot{\boldsymbol{a}}_n \end{bmatrix} = \boldsymbol{g}_{n-1}^{(2)}(\boldsymbol{a}_{n-1}^{(2)}) + \boldsymbol{\rho}_{n-1}^{(2)}$$

$$= \begin{bmatrix} \mathcal{A}_q\left(\mathcal{Q}_a\left(T\dot{\boldsymbol{a}}_{n-1} + \frac{T^2}{2}\ddot{\boldsymbol{a}}_{n-1}\right)\mathcal{Q}_a(\boldsymbol{a}_{n-1})\right) \\ \dot{\boldsymbol{a}}_{n-1} + T\ddot{\boldsymbol{a}}_{n-1} \\ \ddot{\boldsymbol{a}}_{n-1} \end{bmatrix} + \boldsymbol{\rho}_{n-1}^{(1)} \tag{19-28}$$

对于一个恒定角加加速度模型，选取一个12维的状态矢量 $\boldsymbol{a}_n^{(3)} \triangleq [\boldsymbol{a}_n^{\mathrm{T}}, \dot{\boldsymbol{a}}_n^{\mathrm{T}}, \ddot{\boldsymbol{a}}_n^{\mathrm{T}}, \dddot{\boldsymbol{a}}_n^{\mathrm{T}}]^{\mathrm{T}}$，得到过程模型为

$$\boldsymbol{a}_n^{(3)} = \boldsymbol{g}_{n-1}^{(3)}(\boldsymbol{a}_{n-1}^{(3)}) + \boldsymbol{\rho}_{n-1}^{(3)} \tag{19-29}$$

式中：$\boldsymbol{\rho}_{n-1}^{(3)}$ 为这个模型的12维角度噪声，并有

$$\boldsymbol{g}_{n-1}^{(3)}(\boldsymbol{a}_{n-1}^{(3)}) = \begin{bmatrix} \mathcal{A}_q\left(\mathcal{Q}_a\left(T\dot{\boldsymbol{a}}_{n-1}^{\mathrm{T}} + \frac{T^2}{2}\ddot{\boldsymbol{a}}_{n-1}^{\mathrm{T}} + \frac{T^3}{6}\dddot{\boldsymbol{a}}_{n-1}^{\mathrm{T}}\right)\mathcal{Q}_a(\boldsymbol{a}_{n-1})\right) \\ \dot{\boldsymbol{a}}_{n-1}^{\mathrm{T}} + T\ddot{\boldsymbol{a}}_{n-1}^{\mathrm{T}} + \frac{T^2}{2}\dddot{\boldsymbol{a}}_{n-1}^{\mathrm{T}} \\ \ddot{\boldsymbol{a}}_{n-1}^{\mathrm{T}} + T\dddot{\boldsymbol{a}}_{n-1}^{\mathrm{T}} \\ \dddot{\boldsymbol{a}}_{n-1}^{\mathrm{T}} \end{bmatrix} \tag{19-30}$$

19.2.3 组合动态过程模型

平移和定向过程模型可以组合成一个更通用的过程模型。定义组合状态矢

量为

$$\boldsymbol{x}_n^{(i)} = \begin{bmatrix} \boldsymbol{p}_n^{(i)} \\ \boldsymbol{a}_n^{(i)} \end{bmatrix} \tag{19-31}$$

状态噪声模型定义为

$$\boldsymbol{v}_n^{(i)} = \begin{bmatrix} \boldsymbol{v}_n^{(i)} \\ \boldsymbol{\rho}_n^{(i)} \end{bmatrix} \tag{19-32}$$

完整的动态模型可写为

$$\boldsymbol{x}_n^{(i)} = \boldsymbol{f}_{n-1}^{(i)}(\boldsymbol{x}_{n-1}^{(i)}) + \boldsymbol{v}_{n-1}^{(i)} \tag{19-33}$$

$$\boldsymbol{f}_{n-1}^{(i)}(\boldsymbol{x}_{n-1}^{(i)}) = \begin{bmatrix} \boldsymbol{F}^{(i)}\boldsymbol{p}_{n-1}^{(i)} \\ \boldsymbol{g}^{(i)}(\boldsymbol{a}_{n-1}^{(i)}) \end{bmatrix} \tag{19-34}$$

19.2.4 动态过程噪声模型

假定噪声为零均值高斯分布,可以将噪声的形式写为

$$\boldsymbol{v}_n^{(i)} \backsim \mathcal{N}(\boldsymbol{0}_{3(i+3)}, \boldsymbol{Q}^{(i)}) \quad i = 0,1,2,3 \tag{19-35}$$

这里 $\boldsymbol{0}_{3(i+3)}$ 是一个 $3(i+3) \times 3(i+3)$ 的零矩阵。

由文献[3,7]可以发现,恒定位置/方向模型的噪声项代表速度白噪声,其连续过程可简化为

$$\boldsymbol{Q}^{(0)} = \begin{bmatrix} q_t^{(0)} T\boldsymbol{I}_3 & \boldsymbol{0}_3 \\ \boldsymbol{0}_3 & q_r^{(0)} T\boldsymbol{I}_3 \end{bmatrix} \tag{19-36}$$

这里 $q_t^{(0)}$ 和 $q_r^{(0)}$ 分别与平移和旋转速度噪声的标准偏差相关。

对于恒定速度模型,文献[3,7]中给出如下噪声协方差,即

$$\boldsymbol{Q}^{(1)} = \begin{bmatrix} q_t^{(1)} \boldsymbol{Q}_1 & \boldsymbol{0}_6 \\ \boldsymbol{0}_6 & q_r^{(1)} \boldsymbol{Q}_1 \end{bmatrix} \tag{19-37}$$

$$\boldsymbol{Q}_1 = \begin{bmatrix} \frac{T^3}{3}\boldsymbol{I} & \frac{T^2}{2}\boldsymbol{I} \\ \frac{T^2}{2}\boldsymbol{I} & T\boldsymbol{I} \end{bmatrix} \tag{19-38}$$

对于高阶位置方程,可将恒定加速度动态噪声模型[3,7]写为

$$\boldsymbol{Q}^{(2)} = \begin{bmatrix} q_t^{(2)} \boldsymbol{Q}_2 & \boldsymbol{0}_9 \\ \boldsymbol{0}_9 & q_r^{(2)} \boldsymbol{Q}_2 \end{bmatrix} \tag{19-39}$$

$$\boldsymbol{Q}_2 = \begin{bmatrix} \frac{T^5}{20}\boldsymbol{I}_3 & \frac{T^4}{8}\boldsymbol{I}_3 & \frac{T^3}{6}\boldsymbol{I}_3 \\ \frac{T^4}{8}\boldsymbol{I}_3 & \frac{T^3}{3}\boldsymbol{I}_3 & \frac{T^2}{2}\boldsymbol{I}_3 \\ \frac{T^3}{6}\boldsymbol{I}_3 & \frac{T^2}{2}\boldsymbol{I}_3 & T\boldsymbol{I}_3 \end{bmatrix} \tag{19-40}$$

根据文献[8],将恒定加加速度噪声协方差写成

$$\boldsymbol{Q}^{(3)} = \begin{bmatrix} q_t^{(3)}\boldsymbol{Q}_3 & \boldsymbol{0}_{12} \\ \boldsymbol{0}_{12} & q_r^{(3)}\boldsymbol{Q}_3 \end{bmatrix} \tag{19-41}$$

$$\boldsymbol{Q}_3 = \begin{bmatrix} \frac{T^7}{252}\boldsymbol{I}_3 & \frac{T^6}{72}\boldsymbol{I}_3 & \frac{T^5}{30}\boldsymbol{I}_3 & \frac{T^4}{24}\boldsymbol{I}_3 \\ \frac{T^6}{72}\boldsymbol{I}_3 & \frac{T^5}{20}\boldsymbol{I}_3 & \frac{T^4}{8}\boldsymbol{I}_3 & \frac{T^3}{6}\boldsymbol{I}_3 \\ \frac{T^5}{30}\boldsymbol{I}_3 & \frac{T^4}{8}\boldsymbol{I}_3 & \frac{T^3}{3}\boldsymbol{I}_3 & \frac{T^2}{2}\boldsymbol{I}_3 \\ \frac{T^4}{24}\boldsymbol{I}_3 & \frac{T^3}{6}\boldsymbol{I}_3 & \frac{T^2}{2}\boldsymbol{I}_3 & T\boldsymbol{I}_3 \end{bmatrix} \tag{19-42}$$

19.3 观测模型分量

假定有 J 个摄像机采集刚体状态数据,观测刚体表面所标记的特征点。每个特征点相对于物体重心的位置是已知的,由于事先校准,因此公差很小。令 $\boldsymbol{S} = \{\boldsymbol{s}_i = [s_{i,x}, s_{i,y}, s_{i,z}]^{\mathrm{T}}; i = 1, \cdots, N_s\}$ 是表示本体坐标系中特征点的一组固定的三维时间不变位置。因为每个摄像机放置位置不同,所以有不同的视角,每个摄像机能看到标记的不同子集。对于 t_n 时刻的第 j 个摄像机,设可见的特征点的数目为 $K_{j,n}$,将摄像机的每个可见特征点指定为 $k_{1,j,n}, k_{2,j,n}, \cdots, k_{K_{j,n},j,n}$,这样每个 $k_{i,j,n}$ 满足 $\boldsymbol{s}_{k_{i,j,n}} \in \boldsymbol{S}$,其中 $1 \leqslant k_{i,j,n} \leqslant N_s$。假设所有摄像机的图像帧都是同步的,摄像机每 0.005s 同时记录一个图像帧;所有摄像机的位置和方向都是每帧校准,每个摄像机帧标记图像与标记图像位置预测的关联是手动管理。图 19-3 给出的坐标系统描述了特征点投影到摄像机的像平面的原理。

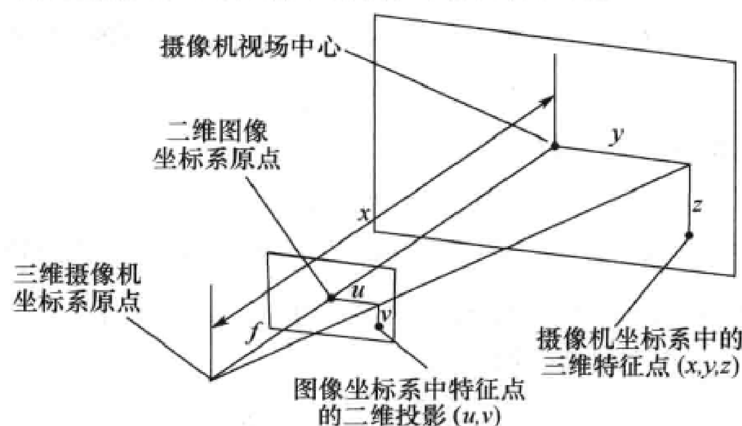


图 19-3 特征点在摄像机的图像平面上的投影

在 t_n 时刻的观测矢量 $\boldsymbol{z}_n$,由 $M_n = \sum_{j=1}^{J} K_{j,n}$ 个二维像素图像标记位置 (u, v) 组成,$\boldsymbol{z}_n$ 可写成

$$z_n = [z_{1,l,u,n}, z_{1,l,v,n}, z_{2,l,u,n}, z_{2,l,v,n}, \cdots, z_{i,j,u,n}, z_{i,j,v,n}, \cdots, z_{K_{J,n},J,u,n}, z_{K_{J,n},J,v,n}]^{\mathrm{T}} \tag{19-43}$$

目前的任务是定义函数 $2M_n$ 维观测矢量分量相对于状态矢量 $\boldsymbol{x}_n$ 的相关性。如果包括观测噪声，观测值可以根据状态矢量写为

$$\boldsymbol{z}_n = \boldsymbol{h}_n(\boldsymbol{x}_n) + \boldsymbol{w}_n \tag{19-44}$$

$$\boldsymbol{h}_n(\boldsymbol{x}_n) = [\boldsymbol{h}_{1,1,n}^{\mathrm{T}}(\boldsymbol{x}_n), \boldsymbol{h}_{2,1,n}^{\mathrm{T}}(\boldsymbol{x}_n), \cdots, \boldsymbol{h}_{i,j,n}^{\mathrm{T}}(\boldsymbol{x}_n), \cdots, \boldsymbol{h}_{K_{J,n}J,n(\boldsymbol{x}_n)}^{\mathrm{T}}]^{\mathrm{T}} \tag{19-45}$$

这样剩下的任务就是定义单个子函数。

假设观测噪声不是零均值高斯过程，为 $\boldsymbol{w}_n \backsim \mathcal{N}(\boldsymbol{0}, \boldsymbol{R})$，其中 $\boldsymbol{R}$ 定义为图像像素对角噪声协方差矩阵的测量方差，可表示为

$$\boldsymbol{R} = \sigma_{pix}^2 \boldsymbol{I}_{2M_n} \tag{19-46}$$

这里 $\boldsymbol{I}_{2M_n}$ 是一个 $2M_n \times 2M_n$ 的单位矩阵。

注意，$\boldsymbol{z}_n$、$\boldsymbol{h}_n$ 和 $\boldsymbol{w}_n$ 都是 $2M_n$ 维，$\boldsymbol{x}_n$ 的维数与动态模型的阶数相关，而 M_n 和 $\boldsymbol{h}_n$ 与时间相关，因为每个摄像机观测到特征点的数目是随时间变化的。为简单起见，假设每个像素分量噪声方差的观测值 $\boldsymbol{\sigma}_{pix}^2$ 与标记特征观测值相同。

附录19. A介绍了点旋转和帧旋转的区别。为将 $\boldsymbol{s}_{k_{i,j,n}}$ 由刚体本体坐标系经过旋转和平移，变换至参考坐标系，应进行点旋转和平移，即

$$\boldsymbol{s}_{k_{i,j,n}}^{(\mathrm{ref})} = \boldsymbol{M}_{\mathrm{ref}} \boldsymbol{s}_{k_{i,j,n}} + \boldsymbol{p}_n \tag{19-47}$$

式中：$\boldsymbol{p}_n$ 为 t_n 时刻位置矢量的质心。旋转矩阵 $\boldsymbol{M}_{\mathrm{ref}}$ 可以通过特征标记点旋转的两个步骤来得。第一步，将旋转 $\boldsymbol{a}_n$ 进行轴角调整，与表示之前旋转的外部四元数 $\boldsymbol{q}_{n-1}$ 相结合。利用式(19－164)定义的函数 $\mathcal{M}_{\mathrm{a}}$ 将轴角矢量 $\boldsymbol{a}_n$ 转换为旋转矩阵 $\boldsymbol{M}_{a_n}$，即

$$\boldsymbol{M}_{a_n} = \mathcal{M}_{\boldsymbol{a}}(\boldsymbol{a}_n) \tag{19-48}$$

利用式(19－163)定义的函数 $\mathcal{M}_{\boldsymbol{q}}$，将外部四元数 $\boldsymbol{q}_{n-1}$ 变换为旋转矩阵 $\boldsymbol{M}_{q_{n-1}}$，即

$$\boldsymbol{M}_{q_{n-1}} = \mathcal{M}_{\boldsymbol{q}}(\boldsymbol{q}_{n-1}) \tag{19-49}$$

将这两个矩阵相乘，得乘积 $\boldsymbol{M}_{\mathrm{ref}}$ 为

$$\boldsymbol{M}_{\mathrm{ref}} = \boldsymbol{M}_{a_n} \boldsymbol{M}_{q_{n-1}} \tag{19-50}$$

将 $\boldsymbol{s}_{k_{i,j,n}}^{(\mathrm{ref})}$ 进行旋转和平移，由参考坐标系变换至第 j 个摄像机的坐标系，分别将摄像机 C_j 的校准位置、方向和焦距定义为 $\boldsymbol{p}_{C_{j,n}}$，$\boldsymbol{q}_{C_{j,n}}$ 和 $f_{C_{j,n}}$，单位为像素，然后进行一个负平移，之后是帧旋转则有

$$\boldsymbol{s}_{k_{i,j,n}}^{(Cj)} = \boldsymbol{M}_{C_{j,n}}^{\mathrm{T}} (\boldsymbol{s}_{k_{i,j,n}}^{(\mathrm{ref})} - \boldsymbol{p}_{C_{j,n}}) \tag{19-51}$$

$$\boldsymbol{M}_{C_j} = \mathcal{M}_{\boldsymbol{q}}(\boldsymbol{q}_{C_j})$$

如图19－3所示，按照下式将三维 $\boldsymbol{s}_{k_{i,j,n}}^{(C_j)}$ 投影到二维图像平面中，则有

$$\boldsymbol{s}_{k_{i,j,n}}^{(\mathrm{image})} = \left(\frac{f_{C_j,n}}{\boldsymbol{s}_{k_{i,j,n},x}^{(C_j)}}\right) \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \boldsymbol{s}_{k_{i,j,n}}^{(C_j)} \tag{19-52}$$

$s_{k_{i,j,n}}^{(\text{image})}$ 的分量$[z_{i,j,u,n}, z_{i,j,v,n}]^{\mathrm{T}} \in z_n$ 可写为

$$z_{i,j,n} = \begin{bmatrix} z_{i,j,u,n} \\ z_{i,j,v,n} \end{bmatrix} = \begin{bmatrix} s_{k_{i,j,n},u}^{(\text{image})} \\ s_{k_{i,j,n},v}^{(\text{image})} \end{bmatrix} \tag{19-53}$$

式中：(u,v) 是图像的像素坐标，以像素为单位。方程(19－53)表示一个摄像机第 i 个特征点图像像素坐标位置。因为观测矢量由 M_N 组 J 个摄像机得的特征点集构成，观测矢量变成

$$z_n = [z_{1,1,n}^{\mathrm{T}}, z_{2,1,n}^{\mathrm{T}}, \cdots, z_{i,j,n}^{\mathrm{T}}, \cdots, z_{K_{J,n},J,n}^{\mathrm{T}}]^{\mathrm{T}} \tag{19-54}$$

其中 $z_{i,j,n} = \boldsymbol{h}_{i,j,n}(\boldsymbol{x}_n)$ 为无噪声的，由式(19－47)～式(19－53)给出。

19.4 估计方法

可以用一些估计方法来进行状态矢量和协方差矩阵的递归估计。下一节将讨论应用于摄影测量法跟踪问题的三种方法，包括 NLLSQ 求解器、高斯 UKF 和无迹组合粒子滤波器，也称为无迹高斯粒子滤波器(UGPF)。本节给出这三种估计方法的应用，第 19.6 节将给出它们的性能比较。

19.4.1 非线性最小平方估计法

以前基于视频图像数据估计刚体的位置和方位的标准方法是 NLSSQ 方法，该方法假定无噪声动态方程，并产生实际测量集和预测测量集之间的最小均方差。在具体实现中，状态矢量中只有刚体在 t_n 时刻的姿势。因为在每个时间步需要重新估计姿势，这里去掉所有时间下标。

由式(19－31)，将 NLSSQ 状态矢量定义为

$$\boldsymbol{x} = [\boldsymbol{p}^{\mathrm{T}}, \boldsymbol{a}^{\mathrm{T}}]^{\mathrm{T}} \tag{19-55}$$

刚体的方向由 $\boldsymbol{a}$ 和表示先验(或初始)方向的四元数 $\boldsymbol{q}$ 相结合进行表示。

式(19－54)给出了观测矢量。观测矢量与状态矢量的关系可以写为

$$z = \boldsymbol{h}(\boldsymbol{x}) + \boldsymbol{w} \tag{19-56}$$

其中 $\boldsymbol{h}(\boldsymbol{x})$ 可按照式(19－45)分解为 $\boldsymbol{h}_{i,j}(\boldsymbol{x})$，每个 $\boldsymbol{h}_{i,j}(\boldsymbol{x})$ 由式(19－47)～式(19－53)中 z 的每个 $z_{i,j}$ 分量给出。假定 $\boldsymbol{w} \backsim \mathcal{N}(\boldsymbol{0}, \boldsymbol{R})$，其中有 $M = \sum_{j=1}^{J} K_j$ 个特征标识的二维图像测量值，$\boldsymbol{R}$ 为 $2M \times 2M$ 对角协方差矩阵，其中对所有的图像像素位置采用同样的方差值。因此可以有

$$\boldsymbol{R} = \sigma_{\text{pix}}^2 \boldsymbol{I}_{2M} \tag{19-57}$$

NLSSQ 方法实质上是对给定观测值 z，通过最小化代价函数，得 $\boldsymbol{x}$ 的加权最小二乘估计，即

$$q(\boldsymbol{x}) = [z - \boldsymbol{h}(\boldsymbol{x})]^{\mathrm{T}} \boldsymbol{R}^{-1} [z - \boldsymbol{h}(\boldsymbol{x})] \tag{19-58}$$

因此，$\boldsymbol{x}$ 的估计值 $\hat{\boldsymbol{x}}$ 可以由下式得，即

$$\nabla_x q(\boldsymbol{x}) = -2[\nabla_x \boldsymbol{h}^{\mathrm{T}}(\boldsymbol{x})]\boldsymbol{R}^{-1}[\boldsymbol{z} - \boldsymbol{h}(\boldsymbol{x})] = \boldsymbol{0} \tag{19-59}$$

将 $\boldsymbol{H}$ 定义为

$$\boldsymbol{H} \triangleq \nabla_x \boldsymbol{h}^{\mathrm{T}}(\boldsymbol{x}) = \nabla_x[\boldsymbol{h}_{1,1}(\boldsymbol{x}), \boldsymbol{h}_{2,1}(\boldsymbol{x}), \cdots, \boldsymbol{h}_{K_J,J}(\boldsymbol{x})]$$

$$= \begin{bmatrix} \dfrac{\partial \boldsymbol{h}_{1,1}}{\partial p_x} & \dfrac{\partial \boldsymbol{h}_{2,1}}{\partial p_x} & \cdots & \dfrac{\partial \boldsymbol{h}_{K_J,J}}{\partial p_x} \\ \dfrac{\partial \boldsymbol{h}_{1,1}}{\partial p_y} & \dfrac{\partial \boldsymbol{h}_{2,1}}{\partial p_y} & \cdots & \dfrac{\partial \boldsymbol{h}_{K_J,J}}{\partial p_y} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial \boldsymbol{h}_{1,1}}{\partial a_z} & \dfrac{\partial \boldsymbol{h}_{2,1}}{\partial a_z} & \cdots & \dfrac{\partial \boldsymbol{h}_{K_J,J}}{\partial a_z} \end{bmatrix} \tag{19-60}$$

非线性函数 $\boldsymbol{h}_{i,j}$ 将 $\boldsymbol{x}$ 变换为第 J 个摄像机图像中第 i 个标记点，进而推导出 $z_{i,j}$。这样式(19 - 59) 可重写为

$$\boldsymbol{H}\boldsymbol{R}^{-1}[\boldsymbol{z} - \boldsymbol{h}(\boldsymbol{x})] = \boldsymbol{0} \tag{19-61}$$

给定状态的某些初始估计(推测)$\hat{\boldsymbol{x}}_0$，将非线性函数 $\boldsymbol{h}(\boldsymbol{x})$ 展开为关于 $\hat{\boldsymbol{x}}_0$ 的泰勒展开式，即

$$\boldsymbol{h}(\boldsymbol{x}) = \boldsymbol{h}(\hat{\boldsymbol{x}}_0) + [\nabla_x \boldsymbol{h}^{\mathrm{T}}(\boldsymbol{x})]_{x=\hat{x}_0}^{\mathrm{T}}[\boldsymbol{x} - \hat{\boldsymbol{x}}_0] + \cdots \tag{19-62}$$

或者，仅保持展开式的线性项

$$\boldsymbol{h}(\boldsymbol{x}) = \boldsymbol{h}(\hat{\boldsymbol{x}}_0) + \boldsymbol{H}_0[\boldsymbol{x} - \hat{\boldsymbol{x}}_0] \tag{19-63}$$

$$\boldsymbol{H}_0 \triangleq [\nabla_x \boldsymbol{h}^{\mathrm{T}}(\boldsymbol{x})]_{x=\hat{x}_0}^{\mathrm{T}} \tag{19-64}$$

方程(19 - 61) 变为

$$\boldsymbol{H}\boldsymbol{R}^{-1}[\boldsymbol{z} - \boldsymbol{h}(\hat{\boldsymbol{x}}_0) - \boldsymbol{H}_0[\boldsymbol{x} - \hat{\boldsymbol{x}}_0]] = \boldsymbol{0} \tag{19-65}$$

令 $\boldsymbol{H} \rightarrow \boldsymbol{H}_0$，求解 $\hat{\boldsymbol{x}}_1 = \boldsymbol{x}$ 得

$$\hat{\boldsymbol{x}}_1 = \hat{\boldsymbol{x}}_0 + [\boldsymbol{H}_0\boldsymbol{R}^{-1}\boldsymbol{H}_0^T]^{-1}\boldsymbol{H}_0\boldsymbol{R}^{-1}[\boldsymbol{z} - \boldsymbol{h}(\hat{\boldsymbol{x}}_0)] \tag{19-66}$$

使用 $\boldsymbol{x}$ 的新估计值作为初始估计并迭代。重复这一过程，在 p 次迭代后，得 Gauss - Newton 标准方程，即

$$\hat{\boldsymbol{x}}_p = \boldsymbol{x}_{p-1} + [\boldsymbol{H}_{p-1}\boldsymbol{R}^{-1}\boldsymbol{H}_{p-1}^{\mathrm{T}}]^{-1}\boldsymbol{H}_{p-1}\boldsymbol{R}^{-1}[\boldsymbol{z} - \boldsymbol{h}(\hat{\boldsymbol{x}}_{p-1})] \tag{19-67}$$

$\hat{\boldsymbol{x}}_p$ 的协方差为

$$\boldsymbol{P}_p = [\boldsymbol{H}_{p-1}\boldsymbol{R}^{-1}\boldsymbol{H}_{p-1}^{\mathrm{T}}]^{-1} \tag{19-68}$$

由式(19 - 57) 可知 $\boldsymbol{R}$ 为对角矩阵，那么上面两个方程可简化为

$$\hat{\boldsymbol{x}}_p = \hat{\boldsymbol{x}}_{p-1} + (\boldsymbol{H}_{p-1}^{\mathrm{T}})^{-1}[\boldsymbol{z} - \boldsymbol{h}(\hat{\boldsymbol{x}}_{p-1})] \tag{19-69}$$

$$\boldsymbol{P}_p = \sigma_{\mathrm{pix}}^{-2}[\boldsymbol{H}_{p-1}\boldsymbol{H}_{p-1}^{\mathrm{T}}]^{-1} \tag{19-70}$$

重复上述过程，直到 $\boldsymbol{e}_p \triangleq \hat{\boldsymbol{x}}_p - \hat{\boldsymbol{x}}_{p-1}$ 的值小于给定的阈值矢量。

因此,在每个 t_n 时间步,可得到一组新的图像 z_n。重复这一计算过程,最终生成估计值 $\hat{\boldsymbol{x}}_n = \hat{\boldsymbol{x}}_{p,n}$,其估计协方差为 $\boldsymbol{P}_{p,n}$。注意,在 Matlab 中,子程序"lsqnonlin"就是 NLSSQ 方法,输入为函数(19 - 58)和变量 $\boldsymbol{x}$,输出为使函数最小化的 $\hat{\boldsymbol{x}}$ 值。尽管这个函数不估计协方差矩阵,但是某些形式的函数会返回雅可比矩阵 $\boldsymbol{H}_{p-1}$,用来计算协方差矩阵估计值。

因为 NLSSQ 公式是迭代但不递归的(在时间上),没有用到动态模型,观测值只与刚体的位置和方向相关,只能估计位置和方向。为了估计任意位置和方向的时间导数,必须用某种形式的平滑滤波器或有限差分滤波器作为后处理器,来得 $\boldsymbol{x}$ 的时间导数。

作为进一步的研究方向,可以考虑在矢量 $\boldsymbol{h}(\boldsymbol{x})$ 中包括一个动态模型或者加入时间上连续的观测值对其进行参数化。这将允许状态矢量中包含导数项,估计过程也可以包含导数项。但这将增加 $\boldsymbol{h}(\boldsymbol{x})$ 和 $\boldsymbol{H}$ 的维数,进而增加计算量。

19.4.2 无迹卡尔曼滤波器法

UKF 是 Sigma 点卡尔曼滤波器的一种,它要求所有的噪声密度都是高斯分布的。对于一个 i 阶模型,其动态方程为非线性的,由式(19 - 33)给出;观测模型也是非线性的,如式(19 - 44)所示。假定动态噪声密度和观测噪声密度均为高斯分布,根据本书第 Ⅱ 部分内容可知,如果对非线性函数进行仿射变换,然后将它们展开为一般多项式,高斯加权矩积分可简化为 Sigma 点集的加权和,这里权值和 Sigma 点取值由所选的 Sigma 点滤波器决定。本节提出用一般 Sigma 点卡尔曼滤波器公式来估计状态矢量,其协方差矩阵基于时间连续图像特征标记观测值集合。

滤波器首先进行状态矢量 $\hat{\boldsymbol{x}}_0 = [\hat{\boldsymbol{p}}_0^{\mathrm{T}}, \hat{\boldsymbol{a}}_0^{\mathrm{T}}]^{\mathrm{T}}$ 的初始估计,得相关的四元数 $\boldsymbol{q}_0$(详见第19.4.4.2节)及其协方差 $\boldsymbol{P}_0^{xx}$。根据第19.2.1节和第19.2.4节,$\boldsymbol{x}$ 的维度、$\boldsymbol{P}^{xx}$ 和 $\boldsymbol{Q}$ 值均取决于所选的模型。

状态 Sigma 点计算公式为

$$\boldsymbol{\chi}_{n-1|n-1}^{(j)} = \hat{\boldsymbol{x}}_{n-1|n-1} + \boldsymbol{D}_{n-1|n-1}\boldsymbol{c}^{(j)} \quad j = 0,1,\cdots,N_s \tag{19 - 71}$$

这里 $\hat{\boldsymbol{x}}_{n-1|n-1} \to \hat{\boldsymbol{x}}_0$ 和 $\boldsymbol{P}_{n-1|n-1}^{xx} \to \boldsymbol{P}_0^{xx}$ 为 Sigma 点的初始集合,$\boldsymbol{D}_{n-1|n-1}$ 定义为

$$\boldsymbol{P}_{n-1|n-1}^{xx} = \boldsymbol{D}_{n-1|n-1}\boldsymbol{D}_{n-1|n-1}^{\mathrm{T}} \tag{19 - 72}$$

式中:N_s 为 Sigma 点的数目;$\boldsymbol{c}^{(j)}$ 为第 j 个 Sigma 点。Sigma 点滤波器中用到的集合 $\{\omega_j, \boldsymbol{c}^{(j)}\}$ 取决于状态矢量的大小和所要实现的 Sigma 点滤波器,可以使用列表 13 - 1 和表 13 - 2 中给出的子程序来生成。

根据第 13 章内容,Sigma 点卡尔曼滤波器的状态预测方程为

$$\boldsymbol{\chi}_{n|n-1}^{(j)} = \boldsymbol{f}(\boldsymbol{\chi}_{n-1|n-1}^{(j)}) \tag{19 - 73}$$

$$\hat{\boldsymbol{x}}_{n|n-1} = \sum_{j=1}^{N_s} w_j \boldsymbol{\chi}_{n-1|n-1}^{(j)} \tag{19-74}$$

$$\boldsymbol{P}_{n|n-1}^{xx} = \sum_{i=1}^{N_s} w_j (\boldsymbol{\chi}_{n|n-1}^{(j)} - \hat{\boldsymbol{x}}_{n|n-1})(\boldsymbol{\chi}_{n|n-1}^{(j)} - \hat{\boldsymbol{x}}_{n|n-1})^{\mathrm{T}} + \boldsymbol{Q} \tag{19-75}$$

观测预测方程为

$$\hat{\boldsymbol{z}}_{n|n-1} = \sum_{j=1}^{N_s} w_j \boldsymbol{h}(\boldsymbol{\chi}_{n|n-1}^{(j)}) \tag{19-76}$$

$$\boldsymbol{P}_{n|n-1}^{zz} = \sum_{j=1}^{N_s} w_j [\boldsymbol{h}(\boldsymbol{\chi}_{n|n-1}^{(j)}) - \hat{\boldsymbol{z}}_{n|n-1}][\boldsymbol{h}(\boldsymbol{\chi}_{n|n-1}^{(j)}) - \hat{\boldsymbol{z}}_{n|n-1}]^{\mathrm{T}} + \boldsymbol{R} \tag{19-77}$$

$$\boldsymbol{P}_{n|n-1}^{xz} = \sum_{j=1}^{N_s} w_j [\boldsymbol{\chi}_{n|n-1}^{(j)} - \hat{\boldsymbol{x}}_{n|n-1}] \cdot [\boldsymbol{h}(\boldsymbol{\chi}_{n|n-1}^{(j)}) - \hat{\boldsymbol{z}}_{n|n-1}]^{\mathrm{T}} \tag{19-78}$$

状态更新方程为

$$\boldsymbol{K}_n \triangleq \boldsymbol{P}_{n|n-1}^{xz} (\boldsymbol{P}_{n|n-1}^{zz})^{-1} \tag{19-79}$$

$$\hat{\boldsymbol{x}}_{n|n} = \hat{\boldsymbol{x}}_{n|n-1} + \boldsymbol{K}_n (\boldsymbol{z}_n^{\mathrm{O}} - \hat{\boldsymbol{z}}_{n|n-1}) \tag{19-80}$$

$$\boldsymbol{P}_{n|n}^{xx} = \boldsymbol{P}_{n|n-1}^{xx} - \boldsymbol{K}_n \boldsymbol{P}_{n|n-1}^{zz} \boldsymbol{K}_n^{\mathrm{T}} \tag{19-81}$$

式中:$\boldsymbol{z}_n^{0}$ 为 t_n 时刻的观测值。一旦在滤波器迭代最后得状态矢量 $\hat{\boldsymbol{x}}_{n|n} = [\hat{\boldsymbol{p}}_{n|n}^{\mathrm{T}}, \hat{\boldsymbol{a}}_{n|n}^{\mathrm{T}}]^{\mathrm{T}}$,由 $\hat{\boldsymbol{a}}_{n|n}$ 和 $\boldsymbol{q}_{n-1}$ 生成新的单位四元数 $\boldsymbol{q}_n$,之后根据下式重置 $\hat{\boldsymbol{a}}_{n|n}$,即

$$\boldsymbol{q}_n = \mathcal{Q}_{\boldsymbol{a}}(\hat{\boldsymbol{a}}_{n|n}) \boldsymbol{q}_{n-1} \tag{19-82}$$

$$\hat{\boldsymbol{a}}_{n|n} \rightarrow [0,0,0]^{\mathrm{T}} \tag{19-83}$$

对于高阶模型和大量特征标记图像尺寸,UKF 是最好的 Sigma 点卡尔曼滤波器,因为所需 Sigma 点的数目为 $N_s = 2n+1$,这里 n 为状态矢量的维数。在 UKF 滤波器的所有应用中,通常选择 $\omega_0 = 0$,因此 $N_s = 2n$ 且 $\omega_j = 1/2n, j = 1, \cdots, 2n$。此外,$\boldsymbol{c}^{(j)} = \sqrt{n\boldsymbol{r}}^{(j)} = \sqrt{n}[1] \in \mathbf{R}^n$(请参阅第 2.2 节的符号解释)。球面单形卡尔曼滤波器(spherical simplex Kalman filter,SSKF)的点更少($N_s = n+1$),简单检验其性能发现,其性能结果与 UKF 相似,这里不详细讨论。

19.4.3 无迹组合粒子滤波器估计法

第 17 章介绍了组合粒子滤波器,即无需进行重采样的粒子滤波方法。这里用它来替代 NLSSQ。组合粒子滤波器的过程框图中利用 UKF 来生成图 17-3 中的重要性采样。在将组合粒子滤波器应用到摄影测量法刚体姿态估计问题时,每个粒子的似然函数 $p(\boldsymbol{z}_n \mid \boldsymbol{x}_n^{(i)})$ 为高斯的,式(19-44)可改写为

$$\boldsymbol{w}_n^{(i)} = \boldsymbol{z}_n^{0} - \boldsymbol{h}_n(\boldsymbol{x}_n^{(i)}) \tag{19-84}$$

因为 $p(\boldsymbol{z}_n \mid \boldsymbol{x}_n^{(i)}) = \mathcal{N}([\boldsymbol{z}_n^{0} - \boldsymbol{h}_n(\boldsymbol{x}_n^{(i)})], \boldsymbol{R})$,可以将粒子似然函数写为

$$p(z_n \mid \boldsymbol{x}_n^{(i)}) = c\exp\left\{-\frac{1}{2}[z_n^{\mathrm{O}} - \boldsymbol{h}_n(\boldsymbol{x}_n^{(i)})]^{\mathrm{T}}\boldsymbol{R}^{-1}[z_n^{\mathrm{O}} - \boldsymbol{h}_n(\boldsymbol{x}_n^{(i)})]\right\} \tag{19-85}$$

除非是数值稳定性需要，不需要包括规范化的 c，因为在权值更新步骤的权值归一化过程中会删除这个变量。在这个应用中，使用20000个粒子。图17-3中已给出整个步骤，这里不再赘述。

19.4.4 估计器初始化

上述所有的估计方法都需要一个初始估计 $\hat{\boldsymbol{x}}$，有一些还需要 $\boldsymbol{P}^{xx}$ 的初始估计。然而，NLLSQ 估计器和所有其他估计器在初始估计所代表的意义上存在一些差异。下面各小节将研究这些问题。

19.4.4.1 NLLSQ 初始化

由于 NLLSQ 是一个迭代滤波器(在一组代码中循环，直到达到容许误差的限制)，在每个时间步骤中必须重新初始化。在投放事件发生之前的时间步中，刚体是静止的，所以初始估计的位置和方向可设为

$$\hat{\boldsymbol{x}}_0 = [\hat{\boldsymbol{p}}_0^{\mathrm{T}}, \hat{\boldsymbol{a}}_0^{\mathrm{T}}]^{\mathrm{T}} = [0,0,0,0,0,0]^{\mathrm{T}} \tag{19-86}$$

这样，$\hat{\boldsymbol{q}}_0$ 的初始估计为一个四元数，表示零角度，即

$$\hat{\boldsymbol{q}}_0 = (1,0,0,0) \tag{19-87}$$

对于其它不同于模拟测试的情况，$\hat{\boldsymbol{p}}_0^{\mathrm{T}}$ 和 $\hat{\boldsymbol{q}}_0$ 额定值可以设为非零值，而 $\hat{\boldsymbol{a}}_0$ 应始终设置为 $\hat{\boldsymbol{a}}_0^{\mathrm{T}} = [0,0,0]^{\mathrm{T}}$。

在投放事件之后，刚体在一个时间步到下一个之间进行平移和旋转，所以从上一时间步估计得的位置和旋转的最终估计可以用来初始化当前时间步的状态矢量。设 $\hat{\boldsymbol{x}}_n^+ = [\hat{\boldsymbol{p}}_n^{+\mathrm{T}}, \hat{\boldsymbol{a}}_n^{+\mathrm{T}}]^{\mathrm{T}}$ 为第 n 帧的后验估计，$\hat{\boldsymbol{q}}_{n-1}$ 表示从第 $n-1$ 的整个旋转，那么 $\hat{\boldsymbol{q}}_n$ 和 $\hat{\boldsymbol{a}}_n$ 可设置为

$$\hat{\boldsymbol{q}}_n = \mathcal{Q}_a(\hat{\boldsymbol{a}}_n^+)\hat{\boldsymbol{q}}_{n-1} \tag{19-88}$$

$$\hat{\boldsymbol{a}}_n = [0,0,0]^{\mathrm{T}} \tag{19-89}$$

第 $n+1$ 帧的先验初始估计可设为 $\hat{\boldsymbol{x}}_{n+1}^- = [\hat{\boldsymbol{p}}_{n+1}^{-\mathrm{T}}, \hat{\boldsymbol{a}}_{n+1}^{-\mathrm{T}}]^{\mathrm{T}}$，有

$$\hat{\boldsymbol{p}}_{n+1}^- = \hat{\boldsymbol{p}}_n^+ \tag{19-90}$$

$$\hat{\boldsymbol{a}}_{n+1}^- = \hat{\boldsymbol{a}}_n^+ = [0,0,0]^{\mathrm{T}} \tag{19-91}$$

19.4.4.2 贝叶斯滤波器初始化

对于除 NLLSQ 外的所有滤波器，需要设置 $\hat{\boldsymbol{x}}_0$、$\boldsymbol{P}_0^{xx}$ 和 $\hat{\boldsymbol{q}}_0$ 的初始值。为保持一

致，这里对所有的滤波器使用相同的初始值。但是，不同的动态模型，所需的初始值也不相同。所以，对于不同阶数的模型，需要不同的初始值：恒定位置／方向，恒定速度，恒加速度以及恒加加速度。因为可以把低阶模型当做其邻近高阶模型的退化，因此仅需初始化最高阶（三阶）的模型变量 $\hat{\boldsymbol{x}}_0^{(3)}$、$\boldsymbol{P}_0^{(3)}$ 和 $\hat{\boldsymbol{q}}_0$，在需要初始化低阶模型时将其截断。因为刚体初始为静止，$\hat{\boldsymbol{x}}_0^{(3)}$ 的初始估计为

$$\hat{\boldsymbol{x}}_0^{(3)} = \begin{bmatrix} \hat{\boldsymbol{p}}_0^{(3)} \\ \hat{\boldsymbol{a}}_0^{(3)} \end{bmatrix} = \boldsymbol{0} \tag{19-92}$$

式中：$\boldsymbol{0}$ 为包含 24 个零的矢量。

模拟测试中 $\hat{\boldsymbol{q}}_0$ 的初始估计为单位四元数，表示角度为零，即

$$\hat{\boldsymbol{q}}_0 = (1,0,0,0) \tag{19-93}$$

在模拟测试以外的其它情况中，$\hat{\boldsymbol{p}}_0$ 和 $\hat{\boldsymbol{q}}_0$ 可以取非零额定值，$\hat{\boldsymbol{a}}_0$ 仍设置为 $\hat{\boldsymbol{a}}_0 = [0,0,0]^{\mathrm{T}}$。因为通常物体都是静止的，$\hat{\boldsymbol{p}}_0$ 和 $\hat{\boldsymbol{a}}_0$ 的导数可仍设置为零。

为初始化 $\boldsymbol{P}_0^{xx}$，令 σ_p、$\sigma_{\dot{p}}$、$\sigma_{\ddot{p}}$ 和 $\sigma_{\dddot{p}}$ 分别为平移位置、速度、加速度和加加速度的标准偏差，σ_a、$\sigma_{\dot{a}}$、$\sigma_{\ddot{a}}$ 和 $\sigma_{\dddot{a}}$ 分别为角度位置、速度、加速度和加加速度的标准偏差。将标准偏差设为

$$\sigma_p = 1 \quad \sigma_{\dot{p}} = 5 \quad \sigma_{\ddot{p}} = 100 \quad \sigma_{\dddot{p}} = 5000 \tag{19-94}$$

$$\sigma_a = 1 \quad \sigma_{\dot{a}} = 1 \quad \sigma_{\ddot{a}} = 50 \quad \sigma_{\dddot{a}} = 2000 \tag{19-95}$$

其中平移的单位是 in 和 s；角度的单位是 deg 和 s。这样，平移协方差子阵的初始值为

$$\boldsymbol{P}_{0,\mathrm{trans}}^{xx(3)} = \begin{bmatrix} \sigma_p^2\boldsymbol{I}_3 & \boldsymbol{0}_3 & \boldsymbol{0}_3 & \boldsymbol{0}_3 \\ \boldsymbol{0}_3 & \sigma_{\dot{p}}^2\boldsymbol{I}_3 & \boldsymbol{0}_3 & \boldsymbol{0}_3 \\ \boldsymbol{0}_3 & \boldsymbol{0}_3 & \sigma_{\ddot{p}}^2\boldsymbol{I}_3 & \boldsymbol{0}_3 \\ \boldsymbol{0}_3 & \boldsymbol{0}_3 & \boldsymbol{0}_3 & \sigma_{\dddot{p}}^2\boldsymbol{I}_3 \end{bmatrix} \tag{19-96}$$

角度协方差子阵的初始值为

$$\boldsymbol{P}_{0,\mathrm{ang}}^{xx(3)} = C_{d2r}\begin{bmatrix} \sigma_a^2\boldsymbol{I}_3 & \boldsymbol{0}_3 & \boldsymbol{0}_3 & \boldsymbol{0}_3 \\ \boldsymbol{0}_3 & \sigma_{\dot{a}}^2\boldsymbol{I}_3 & \boldsymbol{0}_3 & \boldsymbol{0}_3 \\ \boldsymbol{0}_3 & \boldsymbol{0}_3 & \sigma_{\ddot{a}}^2\boldsymbol{I}_3 & \boldsymbol{0}_3 \\ \boldsymbol{0}_3 & \boldsymbol{0}_3 & \boldsymbol{0}_3 & \sigma_{\dddot{a}}^2\boldsymbol{I}_3 \end{bmatrix} \tag{19-97}$$

其中 $C_{d2r} = \pi/180$ 将度变换为弧度。结合这两个子矩阵，构成最初的协方差矩阵，即

$$\boldsymbol{P}_0^{xx(3)} = \begin{bmatrix} \boldsymbol{P}_{0,\mathrm{trans}}^{xx(3)} & \boldsymbol{0}_{12} \\ \boldsymbol{0}_{12} & \boldsymbol{P}_{0,\mathrm{ang}}^{xx(3)} \end{bmatrix} \tag{19-98}$$

在滤波过程开始之后,很少重新初始化贝叶斯估计器,因此仅需进行一次上述初始化过程。

19.5 生成模拟数据

为评估各种跟踪滤波器方法的性能,提出了一种模拟刚体,并生成弹药投放事件的模拟轨迹。这提供了一个模拟的"真实" 轨迹,来创建模拟观测集,例如可采用模拟摄像机通过添加蒙特卡洛集合像素噪声来得到。模拟的摄像机测量值可以用于估计(求解) 方法蒙特卡罗均方根误差分析比较。

19.5.1 模拟刚体特征点

用单位立方体构造一个有8个特征点的刚体,构造一个3×8的矩阵$\boldsymbol{S}$,这里$\boldsymbol{S}$的第i列为特征点$\boldsymbol{s}_i = (s_{i,x}, s_{i,y}, s_{i,z})$,即

$$S = \begin{bmatrix} -1 & 1 & -1 & 1 & -1 & 1 & -1 & 1 \\ -1 & -1 & 1 & 1 & -1 & -1 & 1 & 1 \\ -1 & -1 & -1 & -1 & 1 & 1 & 1 & 1 \end{bmatrix} \tag{19-99}$$

运用一个仿射变换,使立方体变得细长,即

$$\boldsymbol{S} = \boldsymbol{T}\boldsymbol{S} \tag{19-100}$$

$$\boldsymbol{T} = \begin{bmatrix} 200 & 0 & 0 \\ 0 & 30 & 0 \\ 0 & 0 & 30 \end{bmatrix} \tag{19-101}$$

19.5.2 模拟轨迹

为建立一个模拟轨迹,生成一个时间$\boldsymbol{t}$的序列,即

$$\boldsymbol{t} = \{t_n = 0.005 * n, n = 0,1,2,\cdots,200\} \tag{19-102}$$

定义一个阻尼正弦函数(阻尼谐振荡器),即

$$f(t_n,\alpha,\beta,\gamma) = \begin{cases} \alpha e^{-\beta(t_n-t_r)}\{\cos[\gamma(t_n-t_r)]-1\} & t_n \geqslant t_r \\ 0 & t_n < t_r \end{cases} \tag{19-103}$$

式中:t_r为物体投放时间(设$t_r = 0.1$);常数α、β和γ分别为振幅、阻尼系数和频率。在t_r时刻之前物体没有运动,而在t_r时刻之后物体下落。

$f(t_n,\alpha,\beta,\gamma)$ 的前两阶导数为

$$\dot{f}(t_n,\alpha,\beta,\gamma) = \begin{cases} -\alpha e^{-\beta(t_n-t_r)}\{\beta\cos[\gamma(t_n-t_r)] - \\ \beta + \gamma\sin[\gamma(t_n-t_r)]\} & t_n \geqslant t_r \\ 0 & t_n < t_r \end{cases} \tag{19-104}$$

$$\ddot{f}(t_n,\alpha,\beta,\gamma) = \begin{cases} -\alpha e^{-\beta(t_n-t_r)}\{(\beta^2-\gamma^2)\cos[\gamma(t_n-t_r)] - \\ \beta^2 + 2\beta\gamma\sin[\gamma(t_n-t_r)]\} & t_n \geqslant t_r \\ 0 & t_n < t_r \end{cases} \tag{19-105}$$

利用阻尼正弦信号生成 6 个序列，每个对应$(x,y,z,\phi,\theta,\varphi)$ 中六个参数中的一个，其中 ϕ、θ 和 φ 为翻滚、俯仰和偏转的欧拉角，且 $t_n \in \boldsymbol{t}$，则有

$$x(t_n) = f(t_n,\alpha_x,\beta_x,\gamma_x) \quad (\alpha_x,\beta_x,\gamma_x) = (100,2,2) \tag{19-106}$$

$$y(t_n) = f(t_n,\alpha_y,\beta_y,\gamma_y) \quad (\alpha_y,\beta_y,\gamma_y) = (-200,3,1) \tag{19-107}$$

$$z(t_n) = f(t_n,\alpha_z,\beta_z,\gamma_z) \quad (\alpha_z,\beta_z,\gamma_z) = (-2000,1.5,0.5) \tag{19-108}$$

$$\phi(t_n) = f(t_n,\alpha_\phi,\beta_\phi,\gamma_\phi) \quad (\alpha_\phi,\beta_\phi,\gamma_\phi) = (500,1.5,1.25) \tag{19-109}$$

$$\theta(t_n) = f(t_n,\alpha_\theta,\beta_\theta,\gamma_\theta) \quad (\alpha_\theta,\beta_\theta,\gamma_\theta) = (-10,2,10) \tag{19-110}$$

$$\varphi(t_n) = f(t_n,\alpha_\varphi,\beta_\varphi,\gamma_\varphi) \quad (\alpha_\varphi,\beta_\varphi,\gamma_\varphi) = (15,0.8,14) \tag{19-111}$$

然后根据式(19-104)和式(19-105)生成 $\dot{x}(t_n)$、$\dot{y}(t_n)$、$\dot{z}(t_n)$、$\dot{\phi}(t_n)$、$\dot{\theta}(t_n)$、$\dot{\varphi}(t_n)$ 以及 $\ddot{x}(t_n)$、$\ddot{y}(t_n)$、$\ddot{z}(t_n)$、$\ddot{\phi}(t_n)$、$\ddot{\theta}(t_n)$、$\ddot{\varphi}(t_n)$。

可以很容易地构造模拟轨迹的平移分量，即

$$\boldsymbol{p}_{n,\mathrm{true}} = [x(t_n),y(t_n),z(t_n)]^{\mathrm{T}} \tag{19-112}$$

$$\dot{\boldsymbol{p}}_{n,\mathrm{true}} = [\dot{x}(t_n),\dot{y}(t_n),\dot{z}(t_n)]^{\mathrm{T}} \tag{19-113}$$

$$\ddot{\boldsymbol{p}}_{n,\mathrm{true}} = [\ddot{x}(t_n),\ddot{y}(t_n),\ddot{z}(t_n)]^{\mathrm{T}} \tag{19-114}$$

对于旋转分量，必须构造轴角 $\boldsymbol{a}_{n,\mathrm{true}}$ 和单位四元数 $\boldsymbol{q}_{n,\mathrm{true}}$。对于 $\boldsymbol{a}_{n,\mathrm{true}}$，有

$$\boldsymbol{a}_{n,\mathrm{true}} = [0,0,0]^{\mathrm{T}}, \forall n \tag{19-115}$$

为构造 $\boldsymbol{q}_{n,\mathrm{true}}$，首先引入转换函数 $\mathcal{Q}_{\mathrm{Euler}}(\phi,\theta,\varphi):\mathbf{R}^3 \rightarrow \boldsymbol{H}$，将欧拉角$(\phi,\theta,\varphi)$ 转换为单位四元数，即

$$\mathcal{Q}_{\mathrm{Euler}}(\phi,\theta,\varphi) = \boldsymbol{q}_\phi\boldsymbol{q}_\theta\boldsymbol{q}_\varphi \tag{19-116}$$

这里 $\boldsymbol{q}_\phi$、$\boldsymbol{q}_\theta$ 和 $\boldsymbol{q}_\varphi$ 是四元数，代表三个欧拉角，即

$$\boldsymbol{q}_\phi = \left[\cos\frac{\phi}{2},\sin\frac{\phi}{2},0,0\right]^{\mathrm{T}} \tag{19-117}$$

$$\boldsymbol{q}_\theta = \left[\cos\frac{\theta}{2},0,\sin\frac{\theta}{2},0\right]^{\mathrm{T}} \tag{19-118}$$

$$\boldsymbol{q}_\varphi = \left[\cos\frac{\varphi}{2},0,0,\sin\frac{\varphi}{2}\right]^{\mathrm{T}} \tag{19-119}$$

$\boldsymbol{q}_{n,\mathrm{true}}$ 可定义为

$$\begin{aligned}\boldsymbol{q}_{n,\mathrm{true}} &= \mathcal{Q}_{\mathrm{Euler}}(\phi(t_n),\theta(t_n),\varphi(t_n)) \\ &= \boldsymbol{q}_\phi\boldsymbol{q}_\theta\boldsymbol{q}_\varphi\end{aligned} \tag{19-120}$$

为简洁起见，省略了“n,true”下标部分，式(19－15)的四元数乘法得到了间接表述。设

$$\boldsymbol{x}_{n,\mathrm{true}} = [\boldsymbol{p}_{n,\mathrm{true}}^{\mathrm{T}}, \boldsymbol{a}_{n,\mathrm{true}}^{\mathrm{T}}]^{\mathrm{T}} = [\boldsymbol{p}_{n,\mathrm{true}}^{\mathrm{T}}, [0,0,0]^{\mathrm{T}}]^{\mathrm{T}} \tag{19－121}$$

为构造真实数据来比较 $\dot{\boldsymbol{a}}_n$ 和 $\ddot{\boldsymbol{a}}_n$，必须给出模拟的旋转一阶导数和二阶导数表达式。由于使用本体参考坐标系进行求导更为直观，所以这里在本体参考坐标而不是固定参考帧中进行求导运算。将本体参考角度导数记为 $\dot{\boldsymbol{a}}_{n,\mathrm{true}}^{(\mathrm{Body})}$ 和 $\ddot{\boldsymbol{a}}_{n,\mathrm{true}}^{(\mathrm{Body})}$。

回顾著名的四元数微分公式[9]，即

$$\dot{\boldsymbol{q}} = \frac{1}{2}\boldsymbol{\omega}\boldsymbol{q} \tag{19－122}$$

这里 $\boldsymbol{q}$ 为单位四元数，四元数 $\boldsymbol{\omega} = [0,\omega_x,\omega_y,\omega_z]^{\mathrm{T}}$ 可将本体参考旋转速率矢量转换为纯四元数（因为其标量分量为0，因此称其为“纯”）。帧旋转中 $\boldsymbol{q}$、$\dot{\boldsymbol{q}}$ 和 $\boldsymbol{\omega}$ 之间的关系由式(19－122)指定，而 $\boldsymbol{q}_{n,\mathrm{true}}$ 为点旋转。如果 $\boldsymbol{q}$ 为帧旋转，则 $\boldsymbol{q}_{n,\mathrm{true}}$ 等同于点旋转，那么 $\boldsymbol{q}_{n,\mathrm{true}} = \boldsymbol{q}^*$，这里 $\boldsymbol{q}^*$ 为 $\boldsymbol{q}$ 的共轭四元数。即，如果 $\boldsymbol{q} = [q_s,q_x,q_y,q_z]^{\mathrm{T}}$，那么 $\boldsymbol{q}^* = [q_s,-q_x,-q_y,-q_z]^{\mathrm{T}}$。回顾四元数代数的三个事实，对式(19－124)两边取共轭，即：① 如果 $\boldsymbol{w}$ 是纯四元数（其中标量分量 $\omega_s = 0$），那么 $\boldsymbol{w}^* = -\boldsymbol{w}$；② 对于四元数 $\boldsymbol{q}_1$ 和 $\boldsymbol{q}_2$ 的乘积，有 $(\boldsymbol{q}_1\boldsymbol{q}_2)^* = \boldsymbol{q}_2^*\boldsymbol{q}_1^*$；③ 对四元数取两次共轭，得原四元数，即 $(\boldsymbol{q}^*)^* = \boldsymbol{q}$。在式(19－122)中，用 $\boldsymbol{q}$ 替换 $\boldsymbol{q}_{n,\mathrm{true}}^*$，并用 $\boldsymbol{\omega}$ 替换 $-\boldsymbol{\omega}_{n,\mathrm{true}}$，得

$$\dot{\boldsymbol{q}}_{n,\mathrm{true}}^* = -\frac{1}{2}\boldsymbol{\omega}_{n,\mathrm{true}}\boldsymbol{q}_{n,\mathrm{true}}^* \tag{19－123}$$

求解 $\boldsymbol{\omega}_{n,\mathrm{true}}$，得

$$\begin{aligned}\boldsymbol{\omega}_{n,\mathrm{true}} &= -2\dot{\boldsymbol{q}}_{n,\mathrm{true}}^*(\boldsymbol{q}_{n,\mathrm{true}}^*)^{-1} \\ &= -2\dot{\boldsymbol{q}}_{n,\mathrm{true}}^*\frac{(\boldsymbol{q}_{n,\mathrm{true}}^*)^*}{\|\boldsymbol{q}_{n,\mathrm{true}}\|} \\ &= -2\dot{\boldsymbol{q}}_{n,\mathrm{true}}^*\boldsymbol{q}_{n,\mathrm{true}}\end{aligned} \tag{19－124}$$

其中用到了 $\boldsymbol{q}_{n,\mathrm{true}}$ 为单位四元数这一事实。

对式(19－124)等式两边取共轭，得

$$\boldsymbol{\omega}_{n,\mathrm{true}}^* = -\boldsymbol{\omega}_{n,\mathrm{true}} = -2(\dot{\boldsymbol{q}}_{n,\mathrm{true}}^*\boldsymbol{q}_{n,\mathrm{true}})^* = -2\boldsymbol{q}_{n,\mathrm{true}}^*\boldsymbol{q}_{n,\mathrm{true}} \tag{19－125}$$

所以有

$$\boldsymbol{\omega}_{n,\mathrm{true}} = 2\boldsymbol{q}_{n,\mathrm{true}}^*\dot{\boldsymbol{q}}_{n,\mathrm{true}} \tag{19－126}$$

为计算 $\boldsymbol{\omega}_{n,\mathrm{true}}$，必须先求得 $\dot{\boldsymbol{q}}_{n,\mathrm{true}}$。为此，对式(19－120)用乘法规则求微分，即

$$\begin{aligned}\dot{\boldsymbol{q}}_{n,\mathrm{true}} &= \frac{\mathrm{d}}{\mathrm{d}t}(\boldsymbol{q}_\varphi\boldsymbol{q}_\theta\boldsymbol{q}_\phi) \\ &= \dot{\boldsymbol{q}}_\varphi\boldsymbol{q}_\theta\boldsymbol{q}_\phi + \boldsymbol{q}_\varphi\dot{\boldsymbol{q}}_\theta\boldsymbol{q}_\phi + \boldsymbol{q}_\varphi\boldsymbol{q}_\theta\dot{\boldsymbol{q}}_\phi\end{aligned} \tag{19－127}$$

这里 $\dot{\boldsymbol{q}}_\phi$、$\dot{\boldsymbol{q}}_\theta$ 和 $\dot{\boldsymbol{q}}_\varphi$ 可通过对式(19 - 117) ~ 式(19 - 119) 取微分来得,即

$$\dot{\boldsymbol{q}}_\phi = \frac{\dot{\phi}}{2}\left[-\sin\frac{\phi}{2}, \cos\frac{\phi}{2}, 0, 0\right]^{\mathrm{T}} \tag{19 - 128}$$

$$\dot{\boldsymbol{q}}_\theta = \frac{\dot{\theta}}{2}\left[-\sin\frac{\theta}{2}, 0, \cos\frac{\theta}{2}, 0\right]^{\mathrm{T}} \tag{19 - 129}$$

$$\dot{\boldsymbol{q}}_\varphi = \frac{\dot{\varphi}}{2}\left[-\sin\frac{\varphi}{2}, 0, 0, \cos\frac{\varphi}{2}\right]^{\mathrm{T}} \tag{19 - 130}$$

将式(19 - 127) 代入式(19 - 126),得

$$\boldsymbol{\omega}_{n,\mathrm{true}} = 2\boldsymbol{q}^*_{n,\mathrm{true}}(\dot{\boldsymbol{q}}_\varphi \boldsymbol{q}_\theta \boldsymbol{q}_\phi + \boldsymbol{q}_\varphi \dot{\boldsymbol{q}}_\theta \boldsymbol{q}_\phi + \boldsymbol{q}_\varphi \boldsymbol{q}_\theta \dot{\boldsymbol{q}}_\phi) \tag{19 - 131}$$

这样 $\dot{\boldsymbol{a}}^{(\mathrm{Body})}_{n,\mathrm{true}}$ 可写成

$$\dot{\boldsymbol{a}}^{(\mathrm{Body})}_{n,\mathrm{true}} = [\boldsymbol{\omega}_{x,n,\mathrm{true}}, \boldsymbol{\omega}_{y,n,\mathrm{true}}, \boldsymbol{\omega}_{z,n,\mathrm{true}}]^{\mathrm{T}} \tag{19 - 132}$$

用类似的方法,为生成 $\ddot{\boldsymbol{a}}^{(\mathrm{Body})}_{n,\mathrm{true}}$ 必须求得 $\dot{\boldsymbol{\omega}}_{n,\mathrm{true}}$。通过对式(19 - 126) 进行微分计算,即

$$\begin{aligned}\dot{\boldsymbol{\omega}}_{n,\mathrm{true}} &= \frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{\omega}_{n,\mathrm{true}} = \frac{\mathrm{d}}{\mathrm{d}t}\left(2\boldsymbol{q}^*_{n,\mathrm{true}} \frac{d\boldsymbol{q}_{n,\mathrm{true}}}{\mathrm{d}t}\right) \\ &= 2\left[\frac{\mathrm{d}\boldsymbol{q}^*_{n,\mathrm{true}}}{\mathrm{d}t}\frac{\mathrm{d}\boldsymbol{q}_{n,\mathrm{true}}}{\mathrm{d}t} + \boldsymbol{q}^*_{n,\mathrm{true}}\frac{\mathrm{d}^2\boldsymbol{q}_{n,\mathrm{true}}}{\mathrm{d}t^2}\right]\end{aligned} \tag{19 - 133}$$

为计算 $\mathrm{d}\boldsymbol{q}^*_{n,\mathrm{true}}/\mathrm{d}t$,观察到对于所有的 $\boldsymbol{q}$,有 $\mathrm{d}(\boldsymbol{q}\boldsymbol{q}^*)/\mathrm{d}t = 0$,因此

$$\frac{\mathrm{d}}{\mathrm{d}t}(\boldsymbol{q}^*\boldsymbol{q}) = \frac{\mathrm{d}\boldsymbol{q}^*}{\mathrm{d}t}\boldsymbol{q} + \boldsymbol{q}^*\frac{\mathrm{d}\boldsymbol{q}}{\mathrm{d}t} = 0 \tag{19 - 134}$$

进而得

$$\frac{\mathrm{d}\boldsymbol{q}^*}{\mathrm{d}t} = -\boldsymbol{q}^*\frac{\mathrm{d}\boldsymbol{q}}{\mathrm{d}t}\boldsymbol{q}^* = -\boldsymbol{q}^*\dot{\boldsymbol{q}}\boldsymbol{q}^* \tag{19 - 135}$$

将式(19 - 135) 代入式(19 - 133),得

$$\begin{aligned}\dot{\boldsymbol{\omega}}_{n,\mathrm{true}} &= 2[-\boldsymbol{q}^*_{n,\mathrm{true}}\dot{\boldsymbol{q}}_{n,\mathrm{true}}\boldsymbol{q}^*_{n,\mathrm{true}}\dot{\boldsymbol{q}}_{n,\mathrm{true}} + \boldsymbol{q}^*_{n,\mathrm{true}}\ddot{\boldsymbol{q}}_{n,\mathrm{true}}] \\ &= 2\boldsymbol{q}^*_{n,\mathrm{true}}\ddot{\boldsymbol{q}}_{n,\mathrm{true}} - \boldsymbol{\omega}^2_{n,\mathrm{true}}\end{aligned} \tag{19 - 136}$$

对式(19 - 127) 取微分,得 $\ddot{\boldsymbol{q}}_{n,\mathrm{true}}$ 为

$$\begin{aligned}\ddot{\boldsymbol{q}}_{n,\mathrm{true}} = {} & \ddot{\boldsymbol{q}}_\varphi \boldsymbol{q}_\theta \boldsymbol{q}_\phi + \boldsymbol{q}_\varphi \ddot{\boldsymbol{q}}_\theta \boldsymbol{q}_\phi + \boldsymbol{q}_\varphi \boldsymbol{q}_\theta \ddot{\boldsymbol{q}}_\phi + {} \\ & 2\dot{\boldsymbol{q}}_\varphi \dot{\boldsymbol{q}}_\theta \boldsymbol{q}_\phi + 2\dot{\boldsymbol{q}}_\varphi \boldsymbol{q}_\theta \dot{\boldsymbol{q}}_\phi + 2\boldsymbol{q}_\varphi \dot{\boldsymbol{q}}_\theta \dot{\boldsymbol{q}}_\phi\end{aligned} \tag{19 - 137}$$

通过对式(19 - 128) ~ 式(19 - 130) 取微分,得 $\ddot{\boldsymbol{q}}_\phi$、$\ddot{\boldsymbol{q}}_\theta$,和 $\ddot{\boldsymbol{q}}_\varphi$ 为

$$\ddot{\boldsymbol{q}}_\phi = \frac{\ddot{\phi}}{2}\left[-\sin\frac{\phi}{2}, \cos\frac{\phi}{2}, 0, 0\right]^{\mathrm{T}} - \dot{\phi}\boldsymbol{q}_\phi \tag{19 - 138}$$

$$\ddot{\boldsymbol{q}}_\theta = \frac{\ddot{\theta}}{2}\left[-\sin\frac{\theta}{2}, 0, \cos\frac{\theta}{2}, 0\right]^{\mathrm{T}} - \dot{\theta}\boldsymbol{q}_\theta \tag{19 - 139}$$

$$\ddot{\boldsymbol{q}}_{\varphi} = \frac{\ddot{\varphi}}{2}\left[-\sin\frac{\varphi}{2},0,0,\cos\frac{\varphi}{2}\right]^{\mathrm{T}} - \dot{\varphi}\boldsymbol{q}_{\varphi} \tag{19-140}$$

将式(19 - 137)代入式(19 - 136),得

$$\begin{aligned}\dot{\boldsymbol{\omega}}_{n,\text{true}} = 2\boldsymbol{q}^{*}_{n,\text{true}}[(\ddot{\boldsymbol{q}}_{\varphi}\boldsymbol{q}_{\theta}\boldsymbol{q}_{\phi} + \boldsymbol{q}_{\varphi}\ddot{\boldsymbol{q}}_{\theta}\boldsymbol{q}_{\phi} + \boldsymbol{q}_{\varphi}\boldsymbol{q}_{\theta}\ddot{\boldsymbol{q}}_{\phi} + 2\dot{\boldsymbol{q}}_{\varphi}\dot{\boldsymbol{q}}_{\theta}\boldsymbol{q}_{\phi} + \\ 2\dot{\boldsymbol{q}}_{\varphi}\boldsymbol{q}_{\theta}\dot{\boldsymbol{q}}_{\phi} + 2\boldsymbol{q}_{\varphi}\dot{\boldsymbol{q}}_{\theta}\dot{\boldsymbol{q}}_{\phi}) - \boldsymbol{\omega}^{2}_{n,\text{true}}]\end{aligned} \tag{19-141}$$

而 $\ddot{\boldsymbol{a}}^{(\text{Body})}_{n,\text{true}}$ 可写成

$$\ddot{\boldsymbol{a}}^{(\text{Body})}_{n,\text{true}} = [\dot{\omega}_{x,n,\text{true}},\dot{\omega}_{y,n,\text{true}},\dot{\omega}_{z,n,\text{true}}]^{\mathrm{T}} \tag{19-142}$$

19.5.3 模拟摄像机

下面生成模拟数据,假定只有两个模拟摄像机,设置 $J = 2$,并为它们指定位置、方向(单位为°)和焦距(单位为像素),即

$$(x_{C_1},y_{C_1},z_{C_1}) = (490,100,0) \tag{19-143}$$

$$(x_{C_2},y_{C_2},z_{C_2}) = (-490,100,0) \tag{19-144}$$

$$(\phi_{C_1},\theta_{C_1},\varphi_{C_1},f_{C_1}) = (0,0,-168,2500) \tag{19-145}$$

$$(\phi_{C_2},\theta_{C_2},\varphi_{C_2},f_{C_2}) = (0,0,-12,2500) \tag{19-146}$$

然后设置 $\boldsymbol{p}_{C_1}$、$\boldsymbol{q}_{C_1}$、$\boldsymbol{p}_{C_2}$、$\boldsymbol{q}_{C_2}$ 为

$$\boldsymbol{p}_{C_1} = [x_{C_1},y_{C_1},z_{C_1}]^{\mathrm{T}} \tag{19-147}$$

$$\boldsymbol{q}_{C_1} = \mathcal{Q}_{\text{Euler}}(\phi_{C_1},\theta_{C_1},\varphi_{C_1}) \tag{19-148}$$

$$\boldsymbol{p}_{C_2} = [x_{C_2},y_{C_2},z_{C_2}]^{\mathrm{T}} \tag{19-149}$$

$$\boldsymbol{q}_{C_2} = \mathcal{Q}_{\text{Euler}}(\phi_{C_2},\theta_{C_2},\varphi_{C_2}) \tag{19-150}$$

19.5.4 模拟测量

用式(19 - 100)构造模拟刚体之后,对于所有的 i,j 和 n,通过设置 $K_{j,n}$ 和 $k_{i,j,n}$ 来指定在每个时间点上哪个特征点是可见的。为简单起见,整个航迹所有的特征点对两个摄像机都是可见的。很显然,对于 $j = \{1,2\}$,$n = \{0,1,\cdots,200\}$ 有 $K_{j,n} = 8$;而对于 $i = \{1,2,\cdots,8\}$,$j = \{1,2\}$,$n = \{0,1,\cdots,200\}$ 有 $k_{i,j,n} = i$。第 19.3 节已经给出构造测量函数 $\boldsymbol{h}(\boldsymbol{x})$ 所必须的每一项的表达式,然后用 $\boldsymbol{h}(\boldsymbol{x})$ 去生成模拟测量值。对于每个 $n = \{0,1,\cdots,200\}$,通过计算生成测量矢量 $\boldsymbol{z}^{0}_{n,\text{synth}}$,即

$$\boldsymbol{z}^{0}_{n,\text{synth}} = \boldsymbol{h}_{n}(\boldsymbol{x}_{n,\text{true}}) + \boldsymbol{w}_{n,\text{synth}} \tag{19-151}$$

$$\boldsymbol{w}_{n,\text{synth}} \backsim \mathcal{N}(\boldsymbol{0},\boldsymbol{R}_{\text{synth}}) \tag{19-152}$$

$$\boldsymbol{R}_{n,\text{synth}} = \sigma^{2}_{\text{pix}}\boldsymbol{I}_{2M_n} = \sigma^{2}_{\text{pix}}\boldsymbol{I}_{32} \tag{19-153}$$

因为 $M_n = 16 = \sum_{j=1}^{2} K_{j,n}$,$\boldsymbol{I}_{32}$ 为一个 32×32 单位矩阵,而 σ_{pix} 设为 2 个像素。注意,每个时间步都要从 $\mathcal{N}(\boldsymbol{0},\boldsymbol{R}_{\text{synth}})$ 抽取新的蒙特卡洛采样。

19.6 性能对比分析

对第19.4节中给出的所有估计滤波器，用第19.5节描述的模拟测试数据进行试验，以进行性能分析。使用零阶、一阶、二阶和三阶动态模型UKF，与NLLSQ进行性能比较，并与使用二阶动态模型的UGP进行比较。所有情况下使用的性能指标都是均方根(RMS)误差。下面给出的性能结果表明，二阶UKF是最好的滤波器。

图19-4至图19-6给出二阶模型UKF滤波器单次运行的估计轨迹与模拟真实数据比较的结果。观察到图19-5和图19-6中的旋转速率和加速度都是在本体坐标系中的。图19-7给出了横向位置真实航迹一小部分特写，由多个求解器得的估计航迹叠加在真实航迹上。所有估计算法用相同的含噪观测值和初始值作为输入。通过数据分析，得出两个结论：① NLLSQ和零阶动力学模型UKF的估计结果几乎相同。因为零阶模型是一个恒定位置模型，而NLLSQ假定为只有位置的状态矢量，可以预料到NLLSQ和UKF(零阶)在性能上的相似性。② 使用高阶动态模型的滤波器性能优于NLLSQ或使用零阶模型的滤波器，主要是由于减少了估计值相对于真实航迹的变化。第19.5.2节给出的各种滤波器和模型的均方根误差比较得出这一初步的结论。

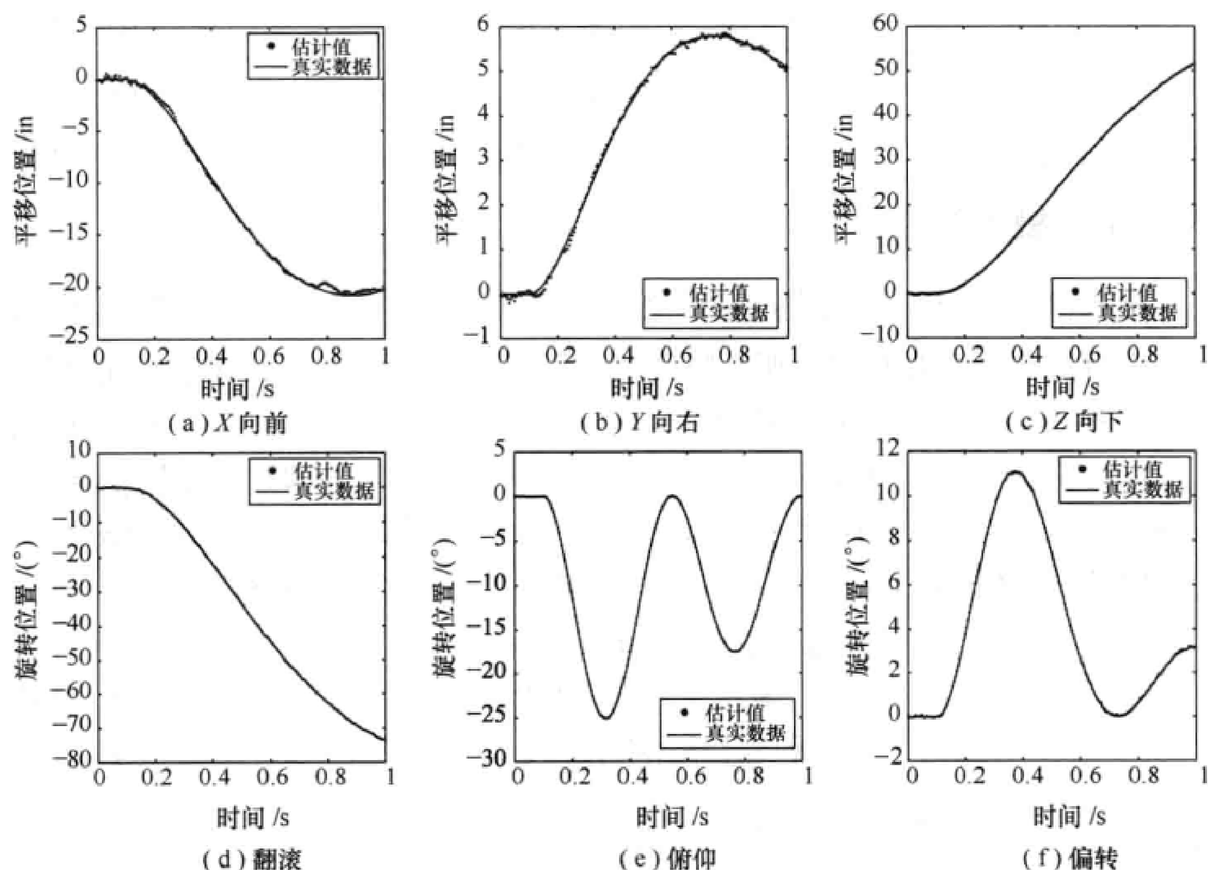


图19-4 用二阶模型UKF滤波器进行平移和旋转的位置

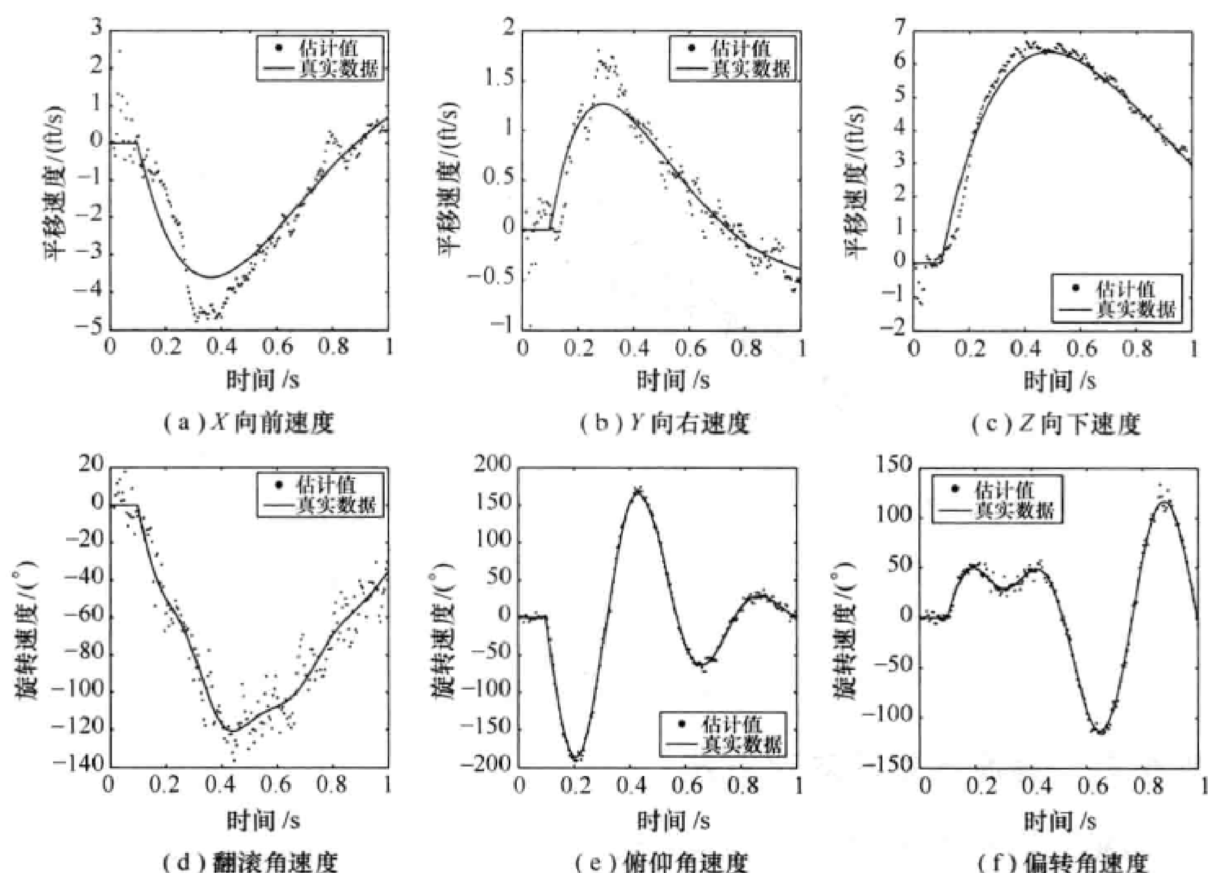


图 19－5　用二阶模型 UKF 滤波器进行平移和旋转的速度

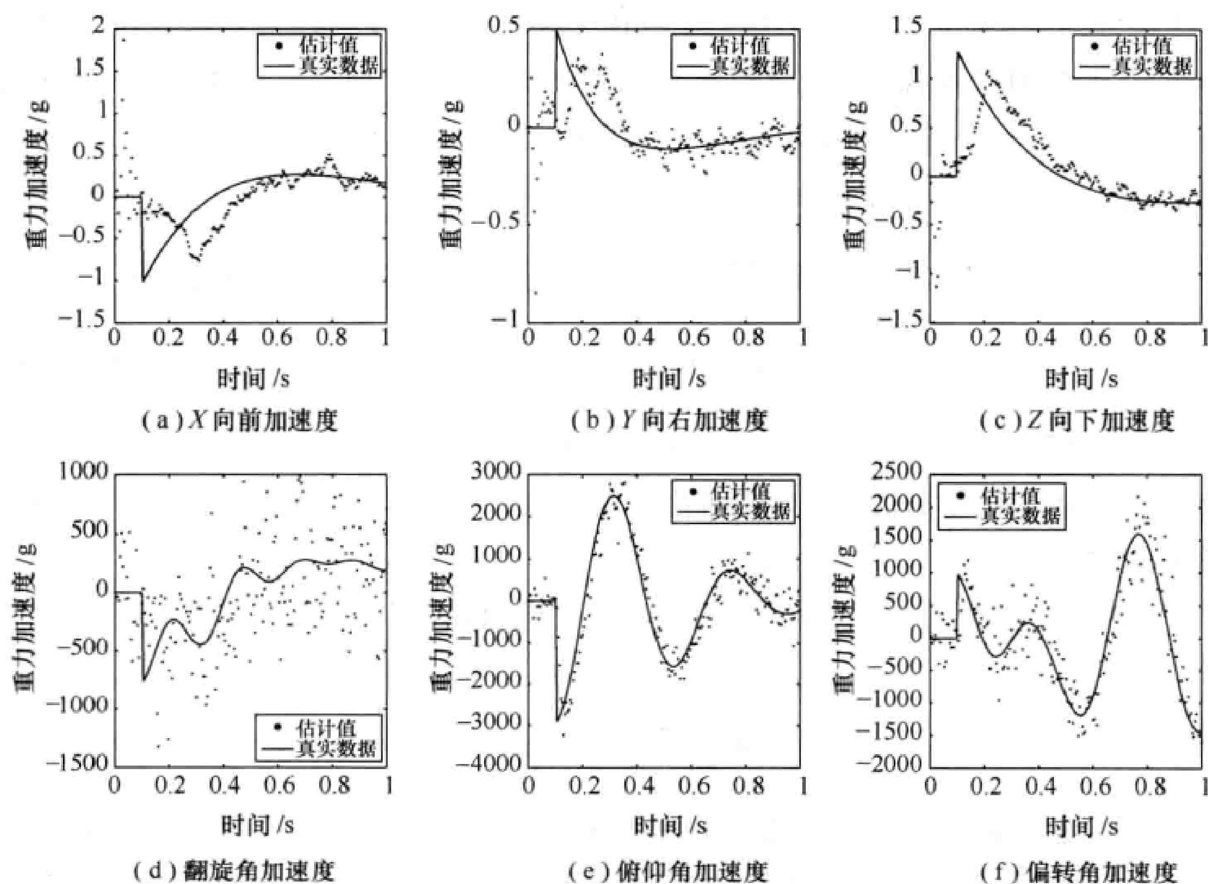


图 19－6　用二阶模型 UKF 滤波器的平移和旋转加速度

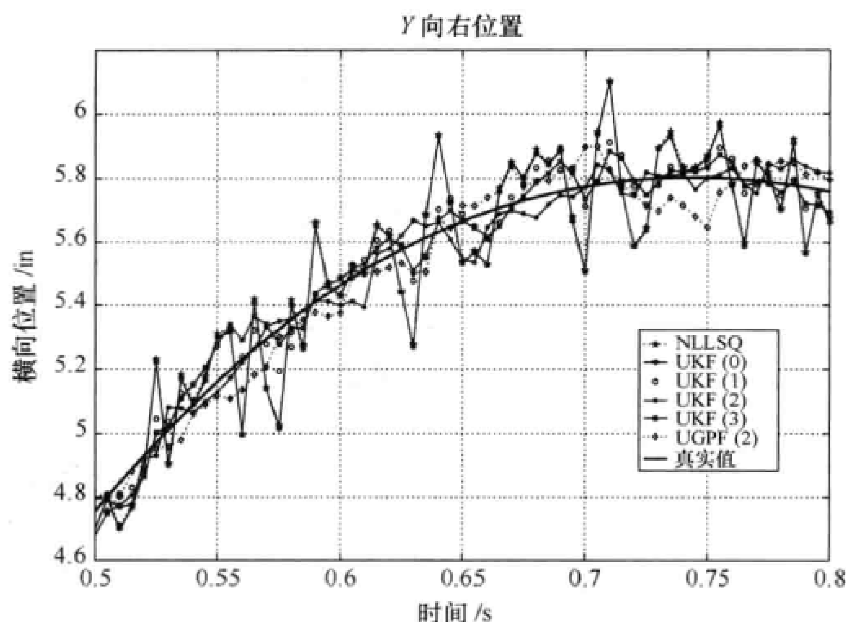


图 19 - 7　用多个求解器使用相同的输入进行横向位置估计

19.6.1　滤波器性能对比方法

性能比较方法用第 19.5.2 节的方法生成一个无噪声的状态矢量航迹$\{\boldsymbol{x}_n, n = 0,1,\cdots,N\}$；然后将这个航迹转换为一组无噪声的模拟真实观测值$\{\boldsymbol{h}_n(\boldsymbol{x}_n), n = 0,1,\cdots,N\}$；最后，生成 $N_{\text{RMS}} = 100$ 组含噪观测值，通过向模拟真实观测值中添加独立零均值高斯噪声，生成含噪观测值的 N_{RMS} 集合$\{\boldsymbol{z}^{0}_{n,r}; n = 0,1,\cdots,N; r = 1,2,\cdots,N_{\text{RMS}}\}$。在每个跟踪滤波器中，用每个含噪观测值集合来产生状态矢量估计的 N_{RMS} 集合$\{\hat{\boldsymbol{x}}_{n|n,r}; n = 0,1,\cdots,N; r = 1,2,\cdots,N_{\text{RMS}}\}$。

第 14 章指出，在 t_n 时刻完整矢量估计 $\hat{\boldsymbol{x}}_{n|n}$ 的协方差矩阵可以写为

$$\boldsymbol{P}^{xx}_{n|n} = \mathcal{E}\{(\boldsymbol{x}_n - \hat{\boldsymbol{x}}_{n|n})(\boldsymbol{x}_n - \hat{\boldsymbol{x}}_{n|n})^{\mathrm{T}}\} \tag{19 - 154}$$

当 $\boldsymbol{x}_n$ 为已知，因为是模拟航迹的情况，而 $\hat{\boldsymbol{x}}_{n|n}$ 是 $\boldsymbol{x}_n$ 的无偏估计，对角项的方根 $\boldsymbol{P}^{xx}_{n|n}$ 可以通过重复滤波器运行来估计，对每次运行使用不同的蒙特卡洛集合含噪观测数据来得条件均方根误差。

例如，滤波器在 t_n 时刻的横向位置(y) 的 RMS 误差为

$$e^{(\text{filter})}_{y,n} = \sqrt{\frac{\sum_{r=1}^{N_{\text{RMS}}}(y_n - \hat{y}^{(\text{filter})}_{n|n,r})^2}{N_{\text{RMS}}}} \tag{19 - 155}$$

比较 6 维问题每个滤波器所得的零阶、一阶、二阶项结果，需要计算并分析 18 个参数的 RMS 误差。为简化分析，使用距离的平方，同时比较位置的 3 个分量，即

$$\| \boldsymbol{p}_n - \boldsymbol{p}_{n|n,r}^{(\text{filter})} \|^2 = (x_n - \hat{x}_{n|n,r}^{(\text{filter})})^2 + (y_n - \hat{y}_{n|n,r}^{(\text{filter})})^2 + (z_n - \hat{z}_{n|n,r}^{(\text{filter})})^2 \tag{19-156}$$

这个结果只有 6 个性能比较，分别是每个 3D 位置及其前两阶导数，还有每个 3D 旋转及其前两阶导数。

3 个平移参数包括 t_n 时刻的位置，RMS 位置误差计算公式为

$$e_{\boldsymbol{p},n}^{(\text{filter})} = \sqrt{\frac{\sum_{r=1}^{N_{\text{RMS}}} \| \boldsymbol{p}_n - \boldsymbol{p}_{n|n,r}^{(\text{filter})} \|^2}{N_{\text{RMS}}}} \tag{19-157}$$

对于位置的一阶和二阶导数，用同样方式计算 $e_{\dot{\boldsymbol{p}},n}^{(\text{filter})}$ 和 $e_{\ddot{\boldsymbol{p}},n}^{(\text{filter})}$。

因为 3D 角度不是矢量，所以不能用同样的方式计算角位移的均方根误差。用单位四元数表示两个方向 $\boldsymbol{q}_1$ 和 $\boldsymbol{q}_2$ 之间的差异，可以有

$$\boldsymbol{q}_\Delta = \boldsymbol{q}_2^* \boldsymbol{q}_1 \tag{19-158}$$

为计算差值，必须将 $\boldsymbol{q}_\Delta$ 转换为轴角表达式，之后才能计算它的范数。因此，可以将某一具体滤波器在 t_n 时刻的 RMS 角误差写成

$$e_{\boldsymbol{a},n}^{(\text{filter})} = \sqrt{\frac{\sum_{r=1}^{N_{\text{RMS}}} \| \mathcal{A}_q(\boldsymbol{q}_n^* \boldsymbol{q}_{n|n,r}^{(\text{filter})} \|^2}{N_{\text{RMS}}}} \tag{19-159}$$

这里 $\mathcal{A}_q: \boldsymbol{H} \to \mathbf{R}^3$（参见式(19-163)）将单位四元数转换为一个轴角矢量，而 $\boldsymbol{q}_{n|n,r}^{(\text{filter})}$ 为表示特定滤波器在 t_n 时刻第 r 次运行的方向估计的四元数，$\boldsymbol{q}_n$ 表示 t_n 时刻的真实方向。

因为旋转速率 $\dot{\boldsymbol{a}}$ 是一个矢量，可以直接比较 $\dot{\boldsymbol{a}}_{n|n,r}^{(\text{filter})}$ 和 $\dot{\boldsymbol{a}}_n$。滤波器估 $\dot{\boldsymbol{a}}$ 计是在参考帧坐标系，而 $\dot{\boldsymbol{a}}$ 的真实航迹是在本体参考坐标系，因此，必须用帧旋转（参见附录 19. A. 3）将 $\dot{\boldsymbol{a}}_{n|n,r}^{(\text{filter})}$ 转换至本体坐标系。这样，特定滤波器在 t_n 时刻的 RMS 角速率误差可写为

$$e_{\dot{\boldsymbol{a}},n}^{(\text{filter})} = \sqrt{\frac{\sum_{r=1}^{N_{\text{RMS}}} \| \boldsymbol{M}_{n,r}^{\text{T}} \dot{\boldsymbol{a}}_{n|n,r}^{(\text{filter})} - \dot{\boldsymbol{a}}_n \|^2}{N_{\text{RMS}}}} \tag{19-160}$$

这里 $\boldsymbol{M}_{n,r} = \mathcal{M}_q(\boldsymbol{q}_{n|n,r}^{(\text{filter})})$ 是一个矩阵，表示特定滤波器在 t_n 时刻第 r 次运行的估计方向，$e_{\ddot{\boldsymbol{a}},n}^{(\text{filter})}$ 可以用同样的方式进行计算。

19.6.2 滤波器对比结果

对于每个滤波器和模型阶数，将 RMS 误差作为时间的函数来进行计算，然后对每种情况生成均值和最大误差（一段时间内）。表 19-1、表 19-2 和表 19-3 分别概括了位置、速度和加速度 RMS 误差的均值和最大值。

表 19－1　模拟数据 RMS 位置误差

滤波器类型	模型阶数	位置/in		方向/(°)	
		均值	极大值	均值	极大值
NLLSQ	0	0.68	0.81	0.22	0.26
UKF	0	0.61	0.72	0.20	0.24
UKF	1	0.30	0.43	0.15	0.22
UKF	2	0.32	0.52	0.15	0.22
UKF	3	0.32	0.47	0.14	0.23
UGPF	2	0.32	0.48	0.15	0.22

表 19－2　模拟数据 RMS 速度误差

滤波器类型	模型阶数	平移速度/(ft/s)		角速度/((°)/s)	
		均值	极大值	均值	极大值
NLLSQ	0	N/A	N/A	N/A	N/A
UKF	0	N/A	N/A	N/A	N/A
UKF	1	0.74	1.43	14.04	22.76
UKF	2	0.71	1.63	12.34	22.18
UKF	3	0.89	1.78	10.56	34.37
UGPF	2	0.72	1.67	12.30	20.39

表 19－3　模拟数据 RMS 加速度误差

滤波器类型	模型阶数	平移加速度/g		角加速度/((°)/s^2)	
		均值	极大值	均值	极大值
NLLSQ	0	N/A	N/A	N/A	N/A
UKF	0	N/A	N/A	N/A	N/A
UKF	1	N/A	N/A	N/A	N/A
UKF	2	0.37	1.70	653.3	3005.9
UKF	3	0.57	1.77	563.9	3085.0
UGPF	2	0.38	1.71	652.6	2990.2

图 19－8、图 19－9 和图 19－10 分别给出每种滤波器/模型次序组合的时程:位置和旋转 RMS 误差、位置和旋转速度 RMS 误差以及位置和旋转加速度 RMS。由表 19－1 ~ 表 19－3 和图 19－8 ~ 图 19－10 明显可见,所有滤波器中使用的模型阶数都大于零,其平均性能在位置上至少提高 50%,在方向上提高 30%。因为物体在 0.1s 时投放,加速度的变化非常大,RMS 误差曲线表明在投放事件后 200ms ~ 300ms 时性能最差。

二阶无迹高斯粒子滤波器(UGPF(2))和二阶无迹卡尔曼滤波器(UKF(2))在 6 个参数类型上的性能几乎相同。因为用高斯随机误差进行模拟

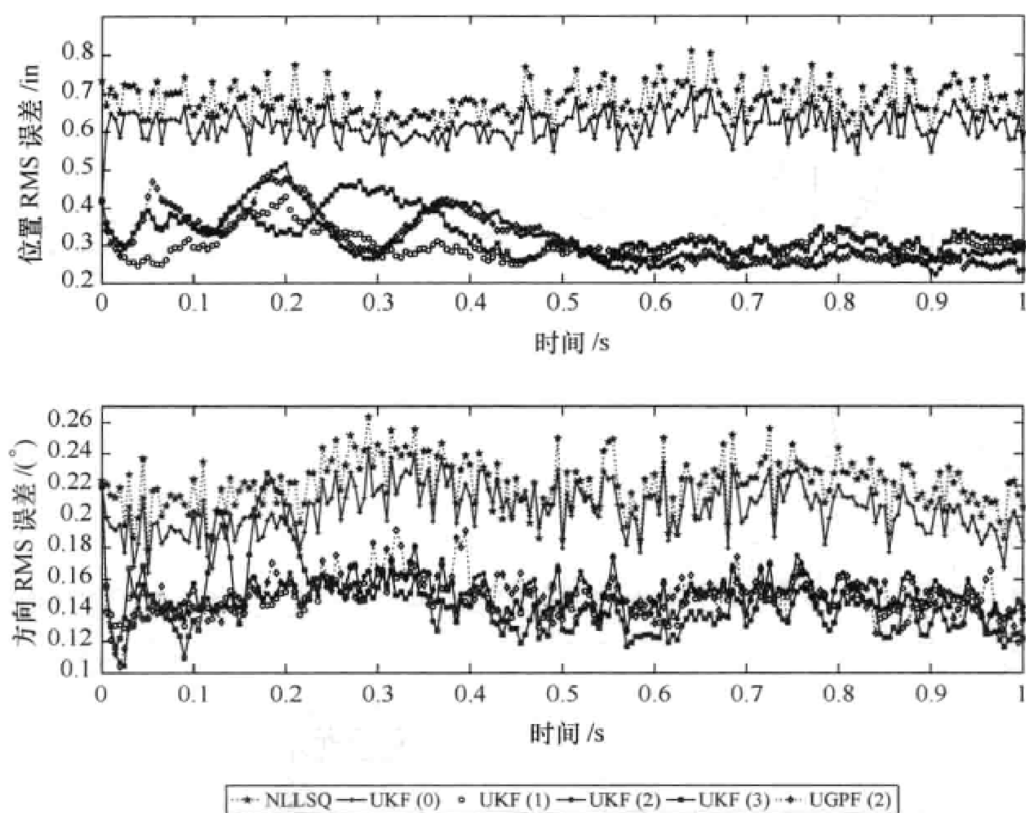


图 19－8　所有跟踪滤波器使用模拟数据得的位置和方向 RMS 误差

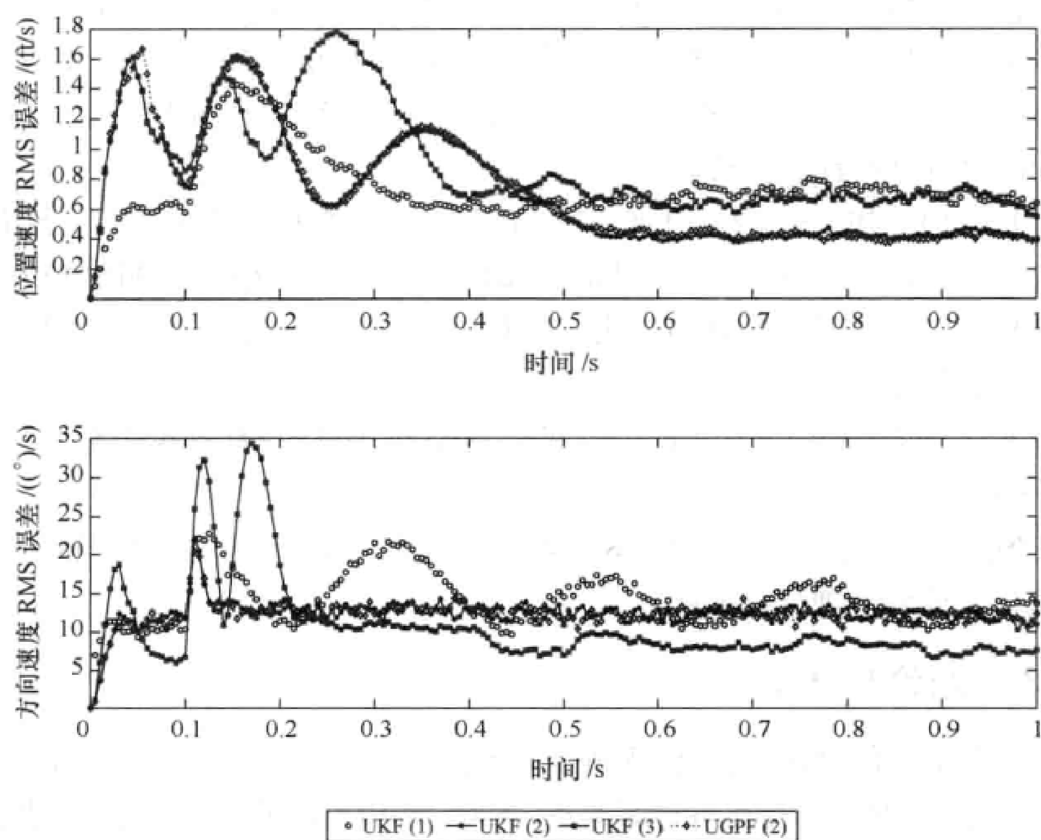


图 19－9　所有跟踪滤波器使用模拟数据得的位置和方向速度的 RMS 误差

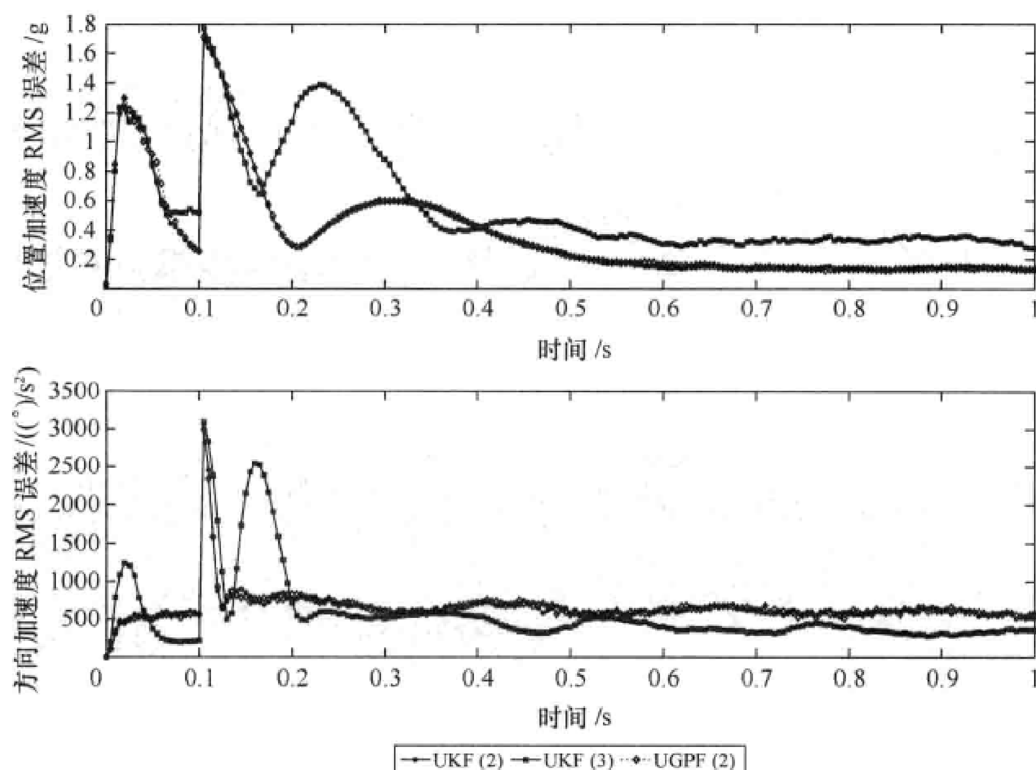


图 19 - 10　所有跟踪滤波器使用模拟数据得的位置和方向加速度的 RMS 误差

测量，UGPF(2) 和 UKF(2) 功能相近，表明动态或测量模型的非线性度不足以拟合状态矢量的概率密度。因为状态矢量密度通常为高斯分布，UGPF(2) 的 RMS 误差与 UKF(2) 的 RMS 误差阶数相同。因此判断，使用粒子滤波器需要相当大的计算量，在这个问题空间里不必要的。如果在实际使用时遇到非高斯测量噪声，可能会重新考虑这个判断。

与 UKF(1) 和 UKF(2) 相比，三阶无迹卡尔曼滤波器（UKF(3)）具有明显的性能问题。尽管与一些参数类型相比，UKF(3) 的均方根误差明显较好，但在某些情况下它的表现比 UKF(1) 或 UKF(2) 更差。当前的三阶动态模型的性能需要进一步研究。

UKF(2) 求解器似乎比 UKF(1) 的性能略好，RMS 统计数据大致为位置、角度和平移速度。

在角速度随时间的变化中，UKF(2) 求解器的 RMS 误差在大约 100ms 稳定，而 UKF(1) 求解器则继续振荡。

19.6.3　结论与展望

在所有求解器测试中，二阶无迹卡尔曼滤波器（UKF(2)）的性能最好，与标准零阶 NLLSQ 求解器相比有显著改进。尽管 UGPF(2) 粒子滤波器的性能与其相似，但计算量非常大，无法实际使用。

在所有阶数大于零的求解器测试中，投放事件的冲击导致加速度跃变后，性

能马上下降。由于这种脉冲的动态模型会显著改善滤波器的性能,动态模型中的自由落体方程不足以表现过程(如受力函数)中的非噪声变化。

解决退化问题的另一种方法是直接测量导致性能下降的加速度。下章将讨论将图像测量与加速度计和陀螺仪的测量结合起来。

拟研究的另一种方法是不仅使用图像中二维特征点的位置,还采取光流技术测量图像序列得二维图像特征点速度。在测量模型中增加特征点图像速度,有助于稳定和提高贝叶斯滤波器三维平移和角速度估计的精度。

附录 19. A 四元数、轴角矢量与旋转

19. A. 1 旋转表示之间的转换

设 $\boldsymbol{a} = [a_x, a_y, a_z]^{\mathrm{T}} \in \mathbf{R}^3$ 为旋转的轴角表示,其中 $\boldsymbol{a}$ 的方向表示旋转轴,$\|\boldsymbol{a}\|$ 表示旋转值(单位为弧度)。

设 $\boldsymbol{q} = [q_s, q_x, q_y, q_z]^{\mathrm{T}} \in \boldsymbol{H}$,$\boldsymbol{H}$ 为四元数空间,是旋转的单位四元数表示,这里 q_s 为实数,q_x、q_y 和 q_z 为纯虚数系数,有 $\sqrt{q_s^2 + q_x^2 + q_y^2 + q_z^2} = 1$。

令 $\boldsymbol{M} \in \mathbf{R}^3 \times \mathbf{R}^3$ 为 3×3 旋转矩阵,这样 $\boldsymbol{M}$ 为正交的,且 $\boldsymbol{M} = +1$。

为将轴角矢量 $\boldsymbol{a}$ 变换为单位四元数 $\boldsymbol{q}$,定义如下函数 $\mathcal{Q}_{\boldsymbol{a}}: \mathbf{R}^3 \to \boldsymbol{H}$,即

$$\boldsymbol{q} = \mathcal{Q}_a(\boldsymbol{a}) = \left[\cos\frac{\|\boldsymbol{a}\|}{2}, \frac{\sin\frac{\|\boldsymbol{a}\|}{2}}{\|\boldsymbol{a}\|}[a_x, a_y, a_z]\right]^{\mathrm{T}} \tag{19-161}$$

为将单位四元数 $\boldsymbol{q}$ 变换为轴角矢量 $\boldsymbol{a}$,定义如下函数 $\mathcal{A}_{\boldsymbol{q}}: \boldsymbol{H} \to \mathbf{R}^3$,即

$$\boldsymbol{a} = \mathcal{A}_q(\boldsymbol{q}) = \frac{2\cos^{-1}}{\sqrt{1 - q_s^2}}[q_x, q_y, q_z]^{\mathrm{T}} \tag{19-162}$$

为将单位四元数 $\boldsymbol{q}$ 变换为旋转矩阵 $\boldsymbol{M}$,定义如下函数 $\mathcal{M}_{\boldsymbol{q}}: \boldsymbol{H} \to \mathbf{R}^3 \times \mathbf{R}^3$,即

$$\boldsymbol{M} = \mathcal{M}_q(\boldsymbol{q}) = \begin{bmatrix} 2(q_s^2 + q_x^2) - 1 & 2(q_x q_y - q_s q_z) & 2(q_x q_z + q_s q_y) \\ 2(q_x q_y + q_s q_z) & 2(q_s^2 + q_y^2) - 1 & 2(q_y q_z - q_s q_x) \\ 2(q_x q_z - q_s q_y) & 2(q_y q_z + q_s q_x) & 2(q_s^2 + q_z^2) - 1 \end{bmatrix} \tag{19-163}$$

为将轴角矢量 $\boldsymbol{a}$ 变换为旋转矩阵 $\boldsymbol{M}$,定义如下函数 $\mathcal{M}_{\boldsymbol{a}}: \mathbf{R}^3 \to \mathbf{R}^3 \times \mathbf{R}^3$,即

$$\boldsymbol{M} = \mathcal{M}_a(\boldsymbol{a}) = \mathcal{M}_q(\mathcal{Q}_a(\boldsymbol{a})) \tag{19-164}$$

为将欧拉角(ϕ, θ, φ)变换为单位四元数 $\boldsymbol{q}$,定义函数 $\mathcal{Q}_{\mathrm{Euler}}(\phi, \theta, \varphi): \mathbf{R}^3$ 为

$$\mathcal{Q}_{\mathrm{Euler}}(\phi, \theta, \varphi) = \boldsymbol{q}_\varphi \boldsymbol{q}_\theta \boldsymbol{q}_\phi \tag{19-165}$$

这里 $\boldsymbol{q}_\phi$、$\boldsymbol{q}_\theta$ 和 $\boldsymbol{q}_\varphi$ 为表示欧拉角的三个四元数,分别为

$$q_\phi = \left[\cos\frac{\phi}{2}, \sin\frac{\phi}{2}, 0, 0\right]^{T} \tag{19-166}$$

$$q_\theta = \left[\cos\frac{\theta}{2}, 0, \sin\frac{\theta}{2}, 0\right]^{T} \tag{19-167}$$

$$q_\varphi = \left[\cos\frac{\varphi}{2}, 0, 0, \sin\frac{\varphi}{2}\right]^{T} \tag{19-168}$$

19. A. 2 方位与旋转的表示方法

用两个量一起表示旋转，一个是状态矢量中的 3D 轴角矢量 $\boldsymbol{a}$，另一个是状态矢量之外的外部单位四元数 $\boldsymbol{q}$，只是在观测模型中用到。在迭代滤波器中，四元数 $\boldsymbol{q}$ 表示此前相对于参考坐标系的刚体方向。矢量 $\boldsymbol{a}$ 表示对 $\boldsymbol{q}$ 的调整，$\boldsymbol{a}$ 只能部分表示当前方向。当旋转可导时，可用 3D 状态矢量来表示 $\dot{\boldsymbol{a}}, \ddot{\boldsymbol{a}}, \dddot{\boldsymbol{a}}, \cdots$。

两个四元数 $\boldsymbol{p}$ 和 $\boldsymbol{q}$ 的乘积结果为四元数 $\boldsymbol{r}$，用下式表示

$$\boldsymbol{r} = \boldsymbol{pq} \tag{19-169}$$

或者

$$\boldsymbol{r} = \begin{bmatrix} r_s \\ r_x \\ r_y \\ r_z \end{bmatrix} = \begin{bmatrix} p_s & -p_x & -p_y & -p_z \\ p_x & p_s & -p_z & p_y \\ p_y & p_z & p_s & -p_x \\ p_z & -p_y & p_x & p_s \end{bmatrix} \begin{bmatrix} q_s \\ q_x \\ q_y \\ q_z \end{bmatrix} \tag{19-170}$$

可以写成矩阵矢量的乘法，即

$$\boldsymbol{r} = \boldsymbol{Pq} \tag{19-171}$$

外部单位四元数在滤波器迭代过程中不会改变，而且只在观测模型中使用。一旦完成滤波器更新步骤，状态矢量 $\boldsymbol{a}_{n|n}$ 中校正后的轴角部分可用于更新外部单位四元数，此后 $\boldsymbol{a}_{n|n}$ 重置为

$$\boldsymbol{q}_n = \mathcal{Q}_a(\boldsymbol{a}_{n|n})\boldsymbol{q}_{n-1} \tag{19-172}$$

$$\boldsymbol{a}_{n|n} \leftarrow [0,0,0]^{T} \tag{19-173}$$

19. A. 3 点旋转与帧旋转

下面给出点旋转和帧旋转的区别。一般来说，在点旋转时，点或矢量旋转而轴的方向保持不变；而在帧旋转时，轴旋转而点或矢量保持不变。为将 3D 矢量 $\boldsymbol{v}$ 旋转单位四元数 $\boldsymbol{q}$ 所指定的角度，用式（19－164）中给出的函数 $\mathcal{M}_q$ 构造一个旋转矩阵，即

$$\boldsymbol{M}_q = \mathcal{M}_q(\boldsymbol{q}) \tag{19-174}$$

进行乘积运算以得旋转矢量 $\boldsymbol{v}^{(\mathrm{rot})}$ 为

$$\boldsymbol{v}^{(\mathrm{rot})} = \boldsymbol{M}_q\boldsymbol{v} \tag{19-175}$$

令 $\boldsymbol{v}^{(f1)}$ 为一个矢量，其元素关于坐标帧 $f1$ 指定，用 $\boldsymbol{q}_{f1\to f2}$ 表示坐标系帧 $f1$ 和

$f2$ 之间的角度差。为实现帧旋转，找到 $\boldsymbol{v}^{(f1)}$ 相对于 $f2$ 的坐标系，即

$$\boldsymbol{M}_{q_{f1\to f2}} = \mathcal{M}_q(\boldsymbol{q}_{f1\to f2}) \tag{19-176}$$

$$\boldsymbol{v}^{(f2)} = \boldsymbol{M}_{q_{f1\to f2}}\boldsymbol{v}^{(f1)} \tag{19-177}$$

注意：帧旋转为点旋转的逆。因为旋转矩阵是正交的，如果 $\boldsymbol{M}$ 为旋转矩阵，那么 $\boldsymbol{M}^{-1} = \boldsymbol{M}^{\mathrm{T}}$。还注意到，$f1$ 和 $f2$ 的原点是一致的。

在观测模型中，用两种方式旋转。① 找到刚体上特征点的坐标，将其经过旋转和平移，变换到参考（飞机）坐标系中物体的估计位置。因为物体上的点被旋转、然后平移到新的位置，这里执行一个点变换，即

$$\boldsymbol{s}^{(\mathrm{ref})} = \boldsymbol{M}_{\mathrm{ref}}\boldsymbol{s} + \boldsymbol{p} \tag{19-178}$$

这里 $\boldsymbol{M}_{\mathrm{ref}}$ 为由估计的物体旋转角度得的旋转矩阵，3D 矢量 $\boldsymbol{p}$ 指定物体重心的平移，而 $\boldsymbol{s}$ 指定3D 特征点在本体坐标系中的位置。② 利用帧变换（负平移后进行一个帧旋转）计算点 $\boldsymbol{s}^{(\mathrm{ref})}$ 相对摄像机坐标系的位置，即

$$\boldsymbol{s}^{(\mathrm{cam})} = \boldsymbol{M}_{\mathrm{cam}}^{\mathrm{T}}(\boldsymbol{s}^{(\mathrm{ref})} - \boldsymbol{p}_{\mathrm{cam}}) \tag{19-179}$$

式中：$\boldsymbol{M}_{\mathrm{cam}}$ 表示摄像机的方向；$\boldsymbol{p}_{\mathrm{cam}}$ 为其位置。

注意：点变换的逆转换为帧变换。所以为了求式（19 - 178）的逆，需要计算

$$\boldsymbol{s} = \boldsymbol{M}_{\mathrm{ref}}^{\mathrm{T}}(\boldsymbol{s}^{(\mathrm{ref})} - \boldsymbol{p}) \tag{19-180}$$

或者为求式（19 - 179）的逆，需要计算

$$\boldsymbol{s}^{(\mathrm{ref})} = \boldsymbol{M}_{\mathrm{cam}}\boldsymbol{s}^{(\mathrm{cam})} + \boldsymbol{p}_{\mathrm{cam}} \tag{19-181}$$

参考文献

[1] Huang TS, Netravali AN. Motion and structure from feature correspondences: a review. Proc. IEEE, 1994, 82(2): 252 - 268.

[2] Iu SL, Wohn K. Estimation of General Rigid Body Motion from a Long Sequence of Images. Department of Computer & Information Science Technical Report, University of Pennsylvania, 1990.

[3] Gennery DB. Visual tracking of known three - dimensional objects. *Int. J. Comput. Vision*, 1992, 7(3): 243 - 270.

[4] Welch GF. SCATT: Incremental Tracking with Incomplete Information. TR96051, Dissertation. Department of Computer Science, UNC - Chapel Hill, 1996.

[5] Welch G, Bishop G. SCATT: incremental tracking with incomplete information. In *Proceedings of the ACM Siggraph* 97. New York: ACM Press, 1997: 333 - 344.

[6] Forseman E, Schug D. Improved analysis techniques for more efficient weapon separation testing. In *Proceedings of the Society of Flight Test Engineers European Chapter*, 2008.

[7] Bar - Shalom Y, Li XR, Kirubarajan T. *Estimation with Applications to Tracking and Navigation*. Wiley, 2001.

[8] Mehrotra K, Mahapatra PR. A jerk model for tracking highly maneuvering targets. *IEEE Trans. Aero. Electron. Syst*, 1997, 33(4): 1094 - 1105.

[9] Online at ae - www. technion. ac. il/admin/serve. php?id = 18558, accessed August 2011.

第 20 章 用摄影测量和惯性测量进行传感器融合

20.1 概　述

第 19 章给出了一种跟踪姿势轨迹(位置和方向的时间函数)方法,仅使用外部的视频摄像传感器估计刚性落体。这种方法中,用安装在刚体表面的多个摄像机采集视频图像。因此,测量值集合仅包含摄像机中特征点的位置,某些方向上刚体由于特征点的遮挡,会有数据的丢失。此外,速度或加速度测量值无法直接获得,所以与位置与方向相关的高阶导数项的估计精度降低。

本章将研究在物体内部额外加装惯性传感器,包括多轴加速度传感器和陀螺仪。这二者合在一起,被称为惯性测量单元(Inertial Measurement Unit,IMU)。IMU 提供沿三个本体坐标轴的线性加速度和相对三个本体坐标轴的角速度,这就需要将 IMU 测量数据时间流与视频数据时间流融合起来,得利用两种异步测量数据流的跟踪滤波器。

在文献[1,2]中首次实现了仅使用单个摄像机(单目成像)得特征点的时间连续图像,在此基础上跟踪刚体的方法。IuandWohn[1] 的论文中,实现了一个动态模型,假设平移加速度和旋转加速度均为常数,而 Lee 等[2] 提出一个动态模型,平移和旋转速度均为常数。文献[3 -6]给出更为先进的方法,使用多个摄像机采集特征点图像,在每个时间步计算刚体的姿态。Gennery[3],Huang 和 Natravali[4] 以及 Ude[5] 均使用非线性最小二乘法(NLLSQ)进行状态矢量估计,假定没有动态噪声。这种方法在时间上不递归,而是在每个时间步进行迭代。Halvorsen 等提出使用时间上递归的扩展卡尔曼滤波器(EKF),来处理多摄像机视频图像特征点缺失的问题。

在文献[8 -12]中,提出用IMU或磁场、角速率和重力传感器(MARG)得的数据估计刚体的方向。在文献[8]中,Algrain 和 Samiie 使用一种加速度计陀螺

线性高斯(AGLG)估计技术,实质上是一种线性最小二乘迭代方法,没有时间递归。文献[9 - 12]中将MARG传感器与EKF相结合,跟踪刚体的方向。这4篇论文对原始算法稍加修改,跟踪人体的手或者肢体的朝向,在时间上递归。

视频和IMU数据的融合是文献[13 - 15]研究的焦点。在所有情况下IMU安装在摄像机的旁边,融合的目的是为了校准摄像机。因为IMU没有安装在跟踪目标的外部,在这些文献提出的方法中,IMU数据不能直接作为测量数据在任何目标状态估计方法中使用。

下面提出一种方法,包括一个安装在刚体上的IMU和外部摄像机以观测刚体上的特征点。由安装在刚体上的IMU得的数据与视频图像数据融合,生成异步数据的单一流,用在跟踪算法的更新步骤中。在这个方法中,采用无迹卡尔曼滤波器(UKF)作为融合跟踪算法,算法的性能结果与单独使用视频图像或IMU数据进行了比较。

图20 - 1给出了一些IMU的图像,独立的或安装在物体上(类似炸弹的刚体)。



图20 - 1　连接发射天线的独立IMU,以及安装在一个500磅物体上的IMU

20.2　刚体运动过程模型

第19.2节描述了刚体运动的动态运动模型,建议读者参考第19章中对这些动态模型的详细描述。第19章的结尾处,试验表明使用二阶恒定加速度可得到相对最好的性能,这里修改状态矢量,使其包括控制平移和旋转运动的速度和加速度矢量。本章将专门分析这个模型,这里的运动模型描述了刚体的平移和旋转运动,是时间的函数。根据瞬时状态矢量 $\boldsymbol{x}(t) = [\boldsymbol{p}^{\mathrm{T}}(t), \dot{\boldsymbol{p}}^{\mathrm{T}}(t), \ddot{\boldsymbol{p}}^{\mathrm{T}}, \boldsymbol{a}^{\mathrm{T}}(t), \dot{\boldsymbol{a}}^{\mathrm{T}}(t), \ddot{\boldsymbol{a}}^{\mathrm{T}}]^{\mathrm{T}}$,根据外部单位四元数 $\boldsymbol{q}(t)$,定义刚体相对于参考笛卡儿帧的质心的平移为 $\boldsymbol{p}(t) = [x(t), y(t), z(t)]^{\mathrm{T}}$,本体相对重心的方向由轴角矢量 $\boldsymbol{a}(t) = [a_x(t), a_y(t), a_z(t)]^{\mathrm{T}}$ 与外部单位四元数 $\boldsymbol{q}(t) = [q_s(t), q_x(t), q_y(t), q_z(t)]^{\mathrm{T}}$ 相结合给出。

一般来说,可以将状态矢量时间转移方程写为

$$x_n = f(x_{n-1}) + v_{n-1} \tag{20-1}$$

这里 v_{n-1} 为导致过程不确定性的噪声,在所有的情况下都假设噪声是高斯加性噪声。

20.3 传感器融合观测模型

定义将状态矢量 x_n 变换至传感器融合观测矢量 z_n 的分量的函数,如果包含观测噪声,此变换可写成

$$z_n = h_n(x_n) + w_n \tag{20-2}$$

这里 $w_n \sim \mathcal{N}(0, R_n)$。观测值由两个独立的分量构成,IMU 分量和摄影测量分量,必须融合在一起来提供观测数据的异步流。

20.3.1 观测模型的惯性测量装置元件

由3个正交加速计和3个正交速率陀螺仪得IMU输出数据,t_n 时刻的IMU观测矢量可写成

$$z_{\mathrm{IMU},n} = [z_{x,n}, z_{y,n}, z_{z,n}, z_{\phi,n}, z_{\theta,n}, z_{\varphi,n}]^{\mathrm{T}} \tag{20-3}$$

这里 $z_{x,n}$,$z_{y,n}$ 和 $z_{z,n}$ 为加速度(用 G 表示),分别用 x - 向前、y - 向右和 z - 向下加速度计进行测量;而 $z_{\phi,n}$,$z_{\theta,n}$ 和 $z_{\varphi,n}$ 为旋转速率(单位为度),分别用翻滚(相对 x 轴逆时针旋转)、俯仰(相对 y 轴逆时针旋转)和偏转(相对 z 轴逆时针旋转)速率陀螺仪测量。注意,因为 IMU 安装到刚体上,当刚体旋转时,IMU 随之进行旋转,这样 $z_{\mathrm{IMU},n}$ 的测量值是关于刚体的本体坐标系,而不是关于参考(飞机)坐标系。

可以根据状态矢量写出 IMU 观测矢量,通过转换方程,即

$$z_{\mathrm{IMU},n} = h_{\mathrm{IMU},n}(x_n) + w_{\mathrm{IMU},n} \tag{20-4}$$

这里 $w_{\mathrm{IMU},n} \sim \mathcal{N}(0, R_{\mathrm{IMU},n})$。将 $R_{\mathrm{IMU},n}$ 定义为由加速度计和速率陀螺仪的测量值方差构成的对角噪声协方差矩阵,即

$$R_{\mathrm{IMU},n} = \begin{bmatrix} \sigma_{\mathrm{acc}}^2 I_3 & 0_3 \\ 0_3 & \sigma_{\mathrm{gyro}}^2 I_3 \end{bmatrix} \tag{20-5}$$

式中:σ_{acc} 和 σ_{gyro} 分别为加速度计标准偏差和角速率测量值标准偏差;I_3 为 3×3 单位矩阵;0_3 为 3×3 零矩阵。

为构造 $h_{\mathrm{IMU},n}(x_n)$,首先将 $g_{\mathrm{ref}} = [g_{x,\mathrm{ref}}, g_{y,\mathrm{ref}}, g_{z,\mathrm{ref}}]^{\mathrm{T}}$ 定义为在参考坐标系的环境加速度(用G表示),地球表面的重力矢量为 $g_{\mathrm{ref}} = [0, 0, -1]^{\mathrm{T}}$。但由于参考的飞机为机动的,并且有荷载 G,g_{ref} 将随时间变化。但是由于飞机比刚体大很多,因此有更大的惯性,这里选取 g_{ref} 为常数,因为与跟踪的刚体相比,它随时间的变化非常缓慢。此外,还忽略由飞机机动性引起的参考坐标系在方向上的改

变。在后续实现中，可以通过考虑这些微小的变化来修改这一过程。还假设 IMU 置于刚体坐标系的中心（通常为其重心），应该加入一个杆臂校正，以弥补加速度计绕刚体重心旋转而引起的向心力和切向线性加速度。

旋转矩阵 $\boldsymbol{M}_{\mathrm{ref}}$ 表示刚体相对参考坐标系的当前方向，可用与第 19 章中观测模型相同的方式构造。利用式(19 - 48) ~ 式(19 - 50)，得

$$\begin{aligned}\boldsymbol{M}_{\mathrm{ref}} &= \boldsymbol{M}_{\boldsymbol{a}_n} \cdot \boldsymbol{M}_{\boldsymbol{q}_{n-1}} \\ \boldsymbol{M}_{\boldsymbol{a}_n} &= \mathcal{M}_{\boldsymbol{a}}(\boldsymbol{a}_n) \\ \boldsymbol{M}_{\boldsymbol{q}_{n-1}} &= \mathcal{M}_{\boldsymbol{q}}(\boldsymbol{q}_{n-1})\end{aligned} \tag{20 - 6}$$

由于 $\boldsymbol{z}_n$ 的测量值是相对于本体坐标系的，可用下式生成无噪本体参考旋转速率 $z_{\phi,n}$，$z_{\theta,n}$ 和 $z_{\varphi,n}$，即

$$\begin{bmatrix} z_{\phi,n} \\ z_{\theta,n} \\ z_{\varphi,n} \end{bmatrix} = \dot{\boldsymbol{a}}_n^{(\mathrm{body})} = \boldsymbol{M}_{\mathrm{ref}}^{\mathrm{T}} \dot{\boldsymbol{a}}_n \tag{20 - 7}$$

用 x、y 和 z 加速度计（在本体—参考坐标系）在 t_n 时刻进行测量，得到3D加速度为

$$\boldsymbol{g}_n^{(\mathrm{body})} = \ddot{\boldsymbol{p}}_n^{(\mathrm{body})} + \boldsymbol{g}_{\mathrm{ref}}^{(\mathrm{body})} \tag{20 - 8}$$

式中：$\ddot{\boldsymbol{p}}_n^{(\mathrm{Body})}$ 是刚体重心的平移加速度；$\boldsymbol{g}_{\mathrm{ref}}^{(\mathrm{Body})}$ 为将参考坐标系旋转至本体坐标系所得的加速度。这样 $\ddot{\boldsymbol{p}}_n^{(\mathrm{Body})} + \boldsymbol{g}_{\mathrm{ref}}^{(\mathrm{Body})}$ 可简单地由下式求得，即

$$\begin{aligned}\ddot{\boldsymbol{p}}_n^{(\mathrm{body})} + \boldsymbol{g}_{\mathrm{ref}}^{(\mathrm{Body})} &= \boldsymbol{M}_{\mathrm{ref}}^{\mathrm{T}} \ddot{\boldsymbol{p}}_n + \boldsymbol{M}_{\mathrm{ref}}^{\mathrm{T}} \boldsymbol{g}_{\mathrm{ref}} \\ &= \boldsymbol{M}_{\mathrm{ref}}^{\mathrm{T}} (\ddot{\boldsymbol{p}}_n + \boldsymbol{g}_{\mathrm{ref}})\end{aligned} \tag{20 - 9}$$

无噪本体参考加速度测量可以表示为

$$\begin{bmatrix} z_{x,n} \\ z_{y,n} \\ z_{z,n} \end{bmatrix} = \boldsymbol{g}_n^{(\mathrm{body})} = \boldsymbol{M}_{\mathrm{ref}}^{\mathrm{T}} (\ddot{\boldsymbol{p}}_n + \boldsymbol{g}_{\mathrm{ref}}) \tag{20 - 10}$$

利用式(20 - 7) ~ 式(20 - 10)，可将 $\boldsymbol{h}_{\mathrm{IMU},n}(\boldsymbol{x}_n)$ 表示为

$$\begin{aligned}\boldsymbol{h}_{\mathrm{IMU},n}(\boldsymbol{x}_n) &= [z_{x,n}, z_{y,n}, z_{z,n}, z_{\phi,n}, z_{\theta,n}, z_{\varphi,n}]^{\mathrm{T}} \\ &= \begin{bmatrix} \boldsymbol{g}_n^{(\mathrm{body})} \\ \dot{\boldsymbol{a}}_n^{(\mathrm{body})} \end{bmatrix}\end{aligned} \tag{20 - 11}$$

20.3.2 观测模型的摄影测量装置

这里简单回顾第 19.3 节中观测模型的定义，来定义观测模型的摄影测量分量。将第 19 章中的分量重新命名为

$$\boldsymbol{z}_n \rightarrow \boldsymbol{z}_{\mathrm{photo},n} \tag{20 - 12}$$

$$\boldsymbol{h}_n(\boldsymbol{x}_n) \rightarrow \boldsymbol{h}_{\mathrm{photo},n}(\boldsymbol{x}_n) \tag{20 - 13}$$

$$\boldsymbol{w}_n \to \boldsymbol{w}_{\text{photo},n} \tag{20-14}$$

$$\boldsymbol{R}_n \to \boldsymbol{R}_{\text{photo},n} \tag{20-15}$$

和第 19 章类似,$\boldsymbol{z}_{\text{photo},n}$、$\boldsymbol{h}_{\text{photo},n}(\boldsymbol{x}_n)$ 和 $\boldsymbol{w}_{\text{photo},n}$ 的维数是 $2M_n$,而 $\boldsymbol{R}_{\text{photo},n}$ 的维数为 $2M_n \times 2M_n$,这里 M_n 为所有摄像机中可见特征点的数目。$\boldsymbol{h}_{\text{photo},n}$ 和 M_n 是时间无关的,因为每个摄像机可观测到的特征点数目可以随时间变化。

20.3.3 组合传感器融合观测模型

为指明 t_n 时刻可获得哪种类型的测量值,定义一个观测标志位 F_n 为

$$F_n = \begin{cases} 1 & \text{只提供 IMU 数据} \\ 2 & \text{只提供摄影测量数据} \\ 3 & \text{提供两种类型数据} \end{cases} \tag{20-16}$$

t_n 时刻的观测矢量 $\boldsymbol{z}_n$ 可写成

$$\boldsymbol{z}_n = \begin{cases} \boldsymbol{z}_{\text{IMU},n} & F_n = 1 \\ \boldsymbol{z}_{\text{photo},n} & F_n = 2 \\ [\boldsymbol{z}_{\text{IMU},n}^{\text{T}}, \boldsymbol{z}_{\text{photo},n}^{\text{T}}]^{\text{T}} & F_n = 3 \end{cases} \tag{20-17}$$

同样,变换函数 $\boldsymbol{h}_n(\boldsymbol{x}_n)$ 可写成

$$\boldsymbol{h}_n(\boldsymbol{x}_n) = \begin{cases} \boldsymbol{h}_{\text{IMU},n}(\boldsymbol{x}_n) & F_n = 1 \\ \boldsymbol{h}_{\text{photo},n}(\boldsymbol{x}_n) & F_n = 2 \\ [\boldsymbol{h}_{\text{IMU},n}^{\text{T}}(\boldsymbol{x}_n), \boldsymbol{h}_{\text{photo},n}^{\text{T}}(\boldsymbol{x}_n)]^{\text{T}} & F_n = 3 \end{cases} \tag{20-18}$$

作为其分量的组合,可将观测噪声假定为零均值高斯过程 $\boldsymbol{w}_n \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{R}_n)$,其中 $\boldsymbol{R}_n$ 定义为对角噪声协方差矩阵,即

$$\boldsymbol{R}_n = \begin{cases} \boldsymbol{R}_{\text{IMU}} & F_n = 1 \\ \boldsymbol{R}_{\text{photo}} & F_n = 2 \\ \begin{bmatrix} \boldsymbol{R}_{\text{IMU}} & \boldsymbol{0}_{\text{SF}} \\ \boldsymbol{0}_{\text{SF}}^{\text{T}} & \boldsymbol{R}_{\text{photo}} \end{bmatrix} & F_n = 3 \end{cases} \tag{20-19}$$

这里 $\boldsymbol{0}_{\text{SF}}$ 为零矩阵的维数 $6 \times 2M_n$。

如果 IMU 时钟与摄影测量时钟不同步,来自 IMU 的数据可能与摄影测量数据异步,连续测量值之间的时间差可能会不一致。因此,在所有跟踪滤波器中必须用时间相关的 T_n,定义为

$$T_n = t_n - t_{n-1} \tag{20-20}$$

20.4 生成模拟数据

为了评估第 19 章中给出各种跟踪方法的性能,提出一种模拟刚体,并生成一个模拟航迹来代表投放弹药的事件。可以由模拟的"真实" 航迹生成由一组虚

拟摄像机看到的模拟含噪观测值的蒙特卡洛集合。蒙特卡洛模拟测量集合被用于各种各样的估计(求解器)方法,用每种估计算法得的跟踪误差均方根(RMS)进行性能的比较分析。

这里重用第 19 章中的模拟航迹和蒙特卡洛含噪摄像机观测值集合,在含噪测量值中添加安装在模拟刚体上的虚拟 IMU 的加速度计和速率陀螺仪的测量值。因此,将进行各种估计方法的比较分析,并分析增加 IMU 测量值得的性能改进。

20.4.1 模拟轨迹

这里使用与第 19.5.2 节相同的模拟航迹,唯一不同的是生成数据速率更快的时间 $\boldsymbol{t}$ 的序列,即

$$\boldsymbol{t} = \{t_n = 0.001 \cdot n \quad n = 0,1,2,\cdots,1000\} \tag{20-21}$$

尽管 IMU 测量值的生成周期更短,为 0.001s,图像测量值仍采用较慢的 0.005s 的周期。为指出这一点,将标记变量 F_n 设置为

$$F_n = \begin{cases} 3 & \{n,\mathrm{mod}5\} = 0 \\ 1 & \text{其他} \end{cases} \tag{20-22}$$

根据式(19-106)~式(19-111)可构造模拟航迹。

20.4.2 模拟摄像机

按照第 19.5.3 节构造两个虚拟摄像机。

20.4.3 模拟测量值

20.4.3.1 IMU 测量值

对于每个 $n = \{0,1,2,\cdots,1000\}$,生成 IMU 模拟测量值矢量 $\boldsymbol{z}^{\mathrm{O}}_{\mathrm{IMU},n,\mathrm{synth}}$,定义为

$$\boldsymbol{z}^{\mathrm{O}}_{\mathrm{IMU},n,\mathrm{synth}} = \boldsymbol{h}_{\mathrm{IMU},n}(\boldsymbol{x}_{n,\mathrm{true}}) + \boldsymbol{w}_{\mathrm{IMU},n,\mathrm{synth}} \tag{20-23}$$

$$\boldsymbol{w}_{\mathrm{IMU},n,\mathrm{synth}} \backsim \mathcal{N}(\boldsymbol{0},\boldsymbol{R}_{\mathrm{IMU},\mathrm{synth}}) \tag{20-24}$$

$$\boldsymbol{R}_{\mathrm{IMU},\mathrm{synth}} = \begin{bmatrix} \sigma^2_{\mathrm{acc}}\boldsymbol{I}_3 & \boldsymbol{0}_3 \\ \boldsymbol{0}_3 & \sigma^2_{\mathrm{gyro}}\boldsymbol{I}_3 \end{bmatrix} \tag{20-25}$$

式中:σ_{acc} 和 σ_{gyro} 分别为 0.1G 和 5.0°/s。

20.4.3.2 摄像机测量值

在时间序列中,取每第 5 个点作为摄像机的测量值。对于每个 $n = \{0,1,2,\cdots,200\}$ 生成相机的模拟测量矢量 $\boldsymbol{z}^{\mathrm{O}}_{\mathrm{IMU},5n,\mathrm{synth}}$,定义为

$$z^0_{\text{photo},5n,\text{synth}} = \boldsymbol{h}_{\text{photo},5n}(\boldsymbol{x}_{5n,\text{true}}) + \boldsymbol{w}_{\text{photo},5n,\text{synth}} \tag{20-26}$$

$$\boldsymbol{w}_{\text{photo},5n,\text{synth}} \backsim \mathcal{N}(\boldsymbol{0}, \boldsymbol{R}_{\text{photo},5n,\text{synth}}) \tag{20-27}$$

$$\boldsymbol{R}_{\text{photo},5n,\text{synth}} = \sigma^2_{\text{pixel}} \boldsymbol{I}_{2M_n} = \sigma^2_{\text{pixel}} \boldsymbol{I}_{32} \tag{20-28}$$

式中:σ_{pixel} 为 2 像素。

20.4.3.3 组合测量值

对于每个 $n = \{0,1,2,\cdots,1000\}$,生成组合模拟测量值矢量 $z^0_{n,\text{synth}}$,定义为

$$z^0_{n,\text{synth}} = \begin{cases} \boldsymbol{h}_{\text{c},n}(\boldsymbol{x}_{n,\text{true}}) + \boldsymbol{w}_{n,\text{synth}} & \{n, \text{mod}5\} = 0 \\ \boldsymbol{h}_{\text{IMU},n}(\boldsymbol{x}_{n,\text{true}}) + \boldsymbol{w}_{n,\text{synth}} & \text{其他} \end{cases} \tag{20-29}$$

$$\boldsymbol{h}_{\text{c},n}(\boldsymbol{x}_{n,\text{true}}) \triangleq [\boldsymbol{h}^{\text{T}}_{\text{IMU},n}(\boldsymbol{x}_{n,\text{true}}), \boldsymbol{h}^{\text{T}}_{\text{photo},n}(\boldsymbol{x}_{n,\text{true}})]^{\text{T}} \tag{20-30}$$

$$\boldsymbol{w}_{n,\text{synth}} \backsim \mathcal{N}(\boldsymbol{0}, \boldsymbol{R}_{n,\text{synth}}) \tag{20-31}$$

$$\boldsymbol{R}_{n,\text{synth}} = \begin{cases} \begin{bmatrix} \boldsymbol{R}_{\text{IMU},\text{synth}} & \boldsymbol{0}_{\text{SF}} \\ \boldsymbol{0}^{\text{T}}_{\text{SF}} & \boldsymbol{R}_{\text{photo},5n,\text{synth}} \end{bmatrix} & \{n, \text{mod}5\} = 0 \\ \boldsymbol{R}_{\text{IMU},\text{synth}} & \text{其他} \end{cases} \tag{20-32}$$

式中:$\boldsymbol{0}_{\text{SF}}$ 为 6×32 维的零矩阵。注意,在每个时间步进行一个新的蒙特卡洛采样 $\boldsymbol{R}_{n,\text{synth}}$。

20.5 估值法

在传感器融合模型比较分析的估计方法包括:仅使用摄影测量数据的方法、仅使用IMU数据的方法以及将IMU数据和视频数据相融合的方法。第19章得出结论,二阶无迹卡尔曼滤波器(UKF(2))在所有测试的求解器当中性能最好,当仅使用摄影测量视频数据时,UKF(2) 相对于标准零阶 NLLSQ 求解器有很大改进。为完整起见,第 20.6 节性能比较分析将包括只使用摄影测量观测数据时 UKF(2) 和 NLLSQ 性能结果。当只用 IMU 数据作为观测值时,用二阶预测 - 校正方法并用 UKF(2) 作为跟踪滤波器。为实现这一点,设 $F_n = 1, \forall n$,这样只有 IMU 数据可用作观测值。

当仅使用摄影测量视频数据时,NLLSQ 和 UKF(2) 估计过程的细节可分别参考第 19.4.1 节和第 19.4.2 节。此处将不再赘述这些方法的细节。第 20.5.1 节将讨论仅使用 IMU 数据时的预测 - 校正估计算法。

20.5.1 惯性测量装置数据初始化

在通用的刚体位置及方向的估计方法中,使用 IMU 采集的平移加速度和角速率测量值,利用初值问题(IVP) 求解器对常微分方程组(ODE) 进行数值求解。有多种可选的 IVP 求解器,这里详细介绍一种简单精确的欧拉梯形或称为预

测 - 校正方法[16,17]的迭代求解器。首先构造一个一阶常微分方程组，与刚体的 IMU 测量值中位置、方向和平移速度相关；然后，简要描述如何对求解器进行初始化，以及如何使用它来求解常微分方程组，以得目标航迹（位置和方向）的时间函数估计。

20.5.1.1 预测 —— 校正 IVP 求解器

通常，可将一阶常微分方程组求解看成一个初值问题。已知初始值 $\boldsymbol{x}(t_0) = \boldsymbol{x}_0$，由矢量函数 $\boldsymbol{f}(t,\boldsymbol{x}(t))$ 来估计未知函数 $\boldsymbol{x}(t)$。$\boldsymbol{x}(t)$ 与其一阶时间导数关系可写成

$$\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} = \boldsymbol{f}(t,\boldsymbol{x}(t)) \tag{20-33}$$

$$\boldsymbol{x}(t_0) = \boldsymbol{x}_0 \tag{20-34}$$

欧拉方法是最简单的求常微分方程积分的正交规则。令 $\boldsymbol{t} = \{t_n, n = 0,1,2,\cdots,N\}$，并指定 $\boldsymbol{x}_n = \boldsymbol{x}(t_n)$。利用欧拉法则，$\boldsymbol{x}_n$ 的预测值可由下式得，即

$$\boldsymbol{x}_n^{\mathrm{pred}} = \boldsymbol{x}_{n-1} + T\boldsymbol{f}(t_{n-1},\boldsymbol{x}_{n-1}) \tag{20-35}$$

这里 $T = t_n - t_{n-1}$。欧拉方法可看做矩形积分规则，仅在区域的左端点对被积函数进行估计。当 $\boldsymbol{f}(t,\boldsymbol{x})$ 为常量时，这一规则是精确的；但当 $\boldsymbol{f}(t,\boldsymbol{x})$ 为线性时，这一规则不精确，其误差与步长 T 的大小成正比。为得准确的估计，需要选取小的步长，但这里无法实现根据预计误差要求以自动的方式来确定步长。为减少误差的不确定性，有必要在欧拉步骤后，在区域的右端点加上第二个函数估计。这种方法称为预测 - 校正方法，以迭代方式使用梯形法则。首先，使用式(20 - 35)，对整个区域执行欧拉步骤，得初值预测；然后，求外推点的函数值 $\boldsymbol{x}_n^{\mathrm{pred}}$，取两个斜率的均值，得 $\boldsymbol{x}_n$ 的第一个校正估计，即

$$\hat{\boldsymbol{x}}_{1,n} = \boldsymbol{x}_{n-1} + \frac{T}{2}[\boldsymbol{f}(t_{n-1},\boldsymbol{x}_{n-1}) + \boldsymbol{f}(t_n,\boldsymbol{x}_n^{\mathrm{pred}})] \tag{20-36}$$

对于 $k > 1$，校正步骤是迭代重复的，$\hat{\boldsymbol{x}}_{k,n}$ 的估计质量用当前估计与上一个估计的比较实现，即

$$e_k = \|\hat{\boldsymbol{x}}_{k,n} - \hat{\boldsymbol{x}}_{k-1,n}\|^2 \tag{20-37}$$

$$\hat{\boldsymbol{x}}_{k,n} = \boldsymbol{x}_{n-1} + \frac{T}{2}[\boldsymbol{f}(t_{n-1},\boldsymbol{x}_{n-1}) + \boldsymbol{f}(t_n,\hat{\boldsymbol{x}}_{k-1,n})] \tag{20-38}$$

当 e_k 与阈值相比足够小时，即 $e_k < \varepsilon$，迭代终止。通常，取 $\varepsilon = \sqrt{\varepsilon_{\mathrm{mach}}}$，其中 $\sqrt{\varepsilon_{\mathrm{mach}}} \simeq 2.2204 \times 10^{-16}$ 是大于零的最小双精度值，如果加 1，则得可区分的大于 1 的数值。当迭代达到阈值，设置 $\boldsymbol{x}_n = \boldsymbol{x}_{k,n}^{\mathrm{corr}}$，然后进行下一步。

20.5.1.2 IMU 问题的常微分方程组

现在构造一个一阶常微分方程组，将本体 - 参考坐标系下 IMU 测量得的旋

转速率与参考坐标系下的旋转角度和速度联系起来，并将本体 - 参考坐标系下的平移加速度与飞机坐标系下的速度和位置联系起来。令 $\boldsymbol{z}^{\mathrm{O}} = [(\boldsymbol{z}_{\alpha}^{\mathrm{O}})^{\mathrm{T}},(\boldsymbol{z}_{\omega}^{\mathrm{O}})^{\mathrm{T}}]^{\mathrm{T}}$ 为由IMU得的测量值，这里 $\boldsymbol{z}_{\alpha}^{\mathrm{O}} = [z_x, z_y, z_z]^{\mathrm{T}}$ 表示3个正交加速度计得的测量值，$\boldsymbol{z}_{\omega}^{\mathrm{O}} = [z_{\phi}, z_{\theta}, z_{\varphi}]^{\mathrm{T}}$ 表示3个正交速率陀螺仪得的测量值。令 $\boldsymbol{x} = [\boldsymbol{p}^{\mathrm{T}}, \dot{\boldsymbol{p}}^{\mathrm{T}}, \boldsymbol{q}^{\mathrm{T}}]^{\mathrm{T}}$ 为待估计的未知状态矢量，其中 $\boldsymbol{p} = [p_x, p_y, p_z]^{\mathrm{T}}$ 表示3维位置，$\dot{\boldsymbol{p}} = [\dot{p}_x, \dot{p}_y, \dot{p}_z]^{\mathrm{T}}$ 表示3维速度，$\boldsymbol{q} = [q_s, q_x, q_y, q_z]^{\mathrm{T}}$ 是一个单位四元数，表示3维方向。注意 $\boldsymbol{z}^{\mathrm{O}}$ 在本体 - 参考坐标系中测量，而 $\boldsymbol{x}$ 在参考坐标系中表示。

为得 $\boldsymbol{q}$ 的常微分方程，从式(19 - 122) 开始，并用 $\boldsymbol{z}_{\omega}^{\mathrm{O}}$ 替换 $\boldsymbol{\omega}$，得

$$\frac{\mathrm{d}\boldsymbol{q}}{\mathrm{d}t} = \frac{1}{2}\boldsymbol{q}\boldsymbol{z}_{\omega}^{\mathrm{O}} \tag{20 - 39}$$

其中四元数乘法 $\boldsymbol{q} \cdot \boldsymbol{z}_{\omega}^{\mathrm{O}}$ 中，被乘数 $\boldsymbol{z}_{\omega}^{\mathrm{O}}$ 用纯四元数 $[0, \boldsymbol{z}_{\omega}^{\mathrm{O}}]^{\mathrm{T}} = [0, z_{\phi}, z_{\theta}, z_{\varphi}]^{\mathrm{T}}$ 来表示。

同样，为得到3维位置 $\boldsymbol{p}$ 与IMU测得的本体 - 参考加速度测量值 $\boldsymbol{z}_{\alpha}^{\mathrm{O}}$ 之间关系，可通过二阶常微分方程得，即

$$\frac{\mathrm{d}^2\boldsymbol{p}}{\mathrm{d}t^2} = \boldsymbol{M}_q\boldsymbol{z}_{\alpha}^{\mathrm{O}} - g \tag{20 - 40}$$

这里旋转矩阵 $\boldsymbol{M}_q = \mathcal{M}_q(\boldsymbol{q})$ 将 $\boldsymbol{z}_{\alpha}^{\mathrm{O}}$ 旋转至参考坐标系，g 为参考坐标系统中的重力加速度。二阶方程式(20 - 40) 可以通过加上一个速度方程，简化为一个一阶方程组，即

$$\frac{\mathrm{d}\boldsymbol{p}}{\mathrm{d}t} = \dot{\boldsymbol{p}} \tag{20 - 41}$$

利用式(20 - 41)，可将式(20 - 40) 重写为

$$\frac{\mathrm{d}\dot{\boldsymbol{p}}}{\mathrm{d}t} = \boldsymbol{M}_q\boldsymbol{z}_{\alpha}^{\mathrm{O}} - g \tag{20 - 42}$$

这样，式(20 - 39)、式(20 - 41) 和式(20 - 42) 构成一个10维的一阶常微分方程组，可充分描述IMU测量值与刚体的位置、方向和平移速度之间的关系。如果定义 $\boldsymbol{x} = [\boldsymbol{p}^{\mathrm{T}}, \dot{\boldsymbol{p}}^{\mathrm{T}}, \boldsymbol{q}^{\mathrm{T}}]^{\mathrm{T}}$，可将式(20 - 33) 写成

$$\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} = \boldsymbol{f}(t, \boldsymbol{x}) = \boldsymbol{f}\left(t, \begin{bmatrix} \boldsymbol{p} \\ \dot{\boldsymbol{p}} \\ \boldsymbol{q} \end{bmatrix}\right) = \begin{bmatrix} \dot{\boldsymbol{p}} \\ \boldsymbol{M}_q\boldsymbol{z}_{\alpha}^{\mathrm{O}} - \boldsymbol{g} \\ \frac{1}{2}\boldsymbol{q}\boldsymbol{z}_{\omega}^{\mathrm{O}} \end{bmatrix} \tag{20 - 43}$$

20.5.1.3 初始化并使用预测 - 校正器对位置和方向进行估计

可通过如下设置对估计模拟刚体航迹的预测 - 校正器进行初始化，即

$$\begin{aligned} \boldsymbol{x}_0 &= [\boldsymbol{p}_0^{\mathrm{T}}, \dot{\boldsymbol{p}}_0^{\mathrm{T}}, \boldsymbol{q}_0^{\mathrm{T}}]^{\mathrm{T}} \\ &= [0,0,0,0,0,0,1,0,0,0]^{\mathrm{T}} \end{aligned} \tag{20 - 44}$$

其中最后四个元素是单位四元数，表示零旋转。注意在其它情况下 $\boldsymbol{x}_0$ 可能包含非零元素，这取决于环境。由于预测 - 校正器是一个非约束的估计器，一旦迭代生成了 t_n 时刻 $\hat{\boldsymbol{x}}_n$ 的估计，四元数的元素 $\hat{\boldsymbol{q}}_n$ 不一定是单位的，因此不能正确地表示方向。

为校正这一问题，在进行下一次迭代之前对 $\hat{\boldsymbol{q}}_n$ 进行规范化，将其规范至至单位长度，即

$$\hat{\boldsymbol{q}}_n = \frac{\hat{\boldsymbol{q}}_n}{\|\hat{\boldsymbol{q}}_n\|} \tag{20-45}$$

20.6 性能对比分析

用第20.4节给出的模拟试验数据，进行以下性能分析。注意，和第19.6节用到的模拟航迹是相同的（图像数据具有相同的噪声）。这里对传感器融合估计滤波器（UKF(2)）的性能进行比较，对两种仅使用摄影测量数据的估计器（NLLSQ 和 UKF(2)）以及两种仅使用IMU数据的估计器（预测 - 校正器和UKF(2)）进行对比。在所有情况下所使用的性能指标都是均方根（RMS）误差。下面将会给出，传感器融合估计器的均方根误差最小，因此是最好的估计器。

图20 - 2 ~ 图20 - 4 给出一次运行的估计轨迹分量中，传感器融合滤波器与真实跟踪分量的对比。图20 - 5 给出了横向位置真实航迹一小部分的特写，由多个估计器得的估计航迹叠加在真实航迹上。所有估计算法用相同的含噪观测值和初始值作为输入。通过图表分析，可得到以下观察结果。

仅使用IMU数据情况下，两种估计方法都是不准确的，尤其是UKF。众所周知，仅使用IMU数据的估计器对初始条件和校准偏差很敏感，但在这种（模拟）情况下，初始条件和校准是精确的。而且，由于观测值随机噪声漂移，使得估计不准确，尤其当求解器的估计值一个太大而另一个很小时。

仅使用摄影测量数据时，两种估计都是准确的，但并不精确。UKF(2) 似乎比 NLLSQ 更为准确（也更精确），主要是由于其状态矢量的估计中包含导数项。得出以下结论：在所有测试的估计方法中，传感器融合的结果是最准确的，因为仅使用摄影测量的估计器测量空间参数，对于导数只能估计。而仅使用IMU的估计器测量导数，必须估计空间参数，融合估计器，以空间和导数测量值为基础进行估计。

由图20 - 5 可很容易观察到传感器融合估计器的性能，将其航迹与仅使用摄影测量法的UKF相比较，用IMU数据采样率的1/5进行采样。在摄影测量观测值到达前，传感器融合的趋势很平稳，而在每个第五采样点，向一个或另一个方向移动。它沿着与仅使用摄影测量的UKF相同的方向移动，但由于IMU数据提供了更准确的导数估计，移动数量更小。

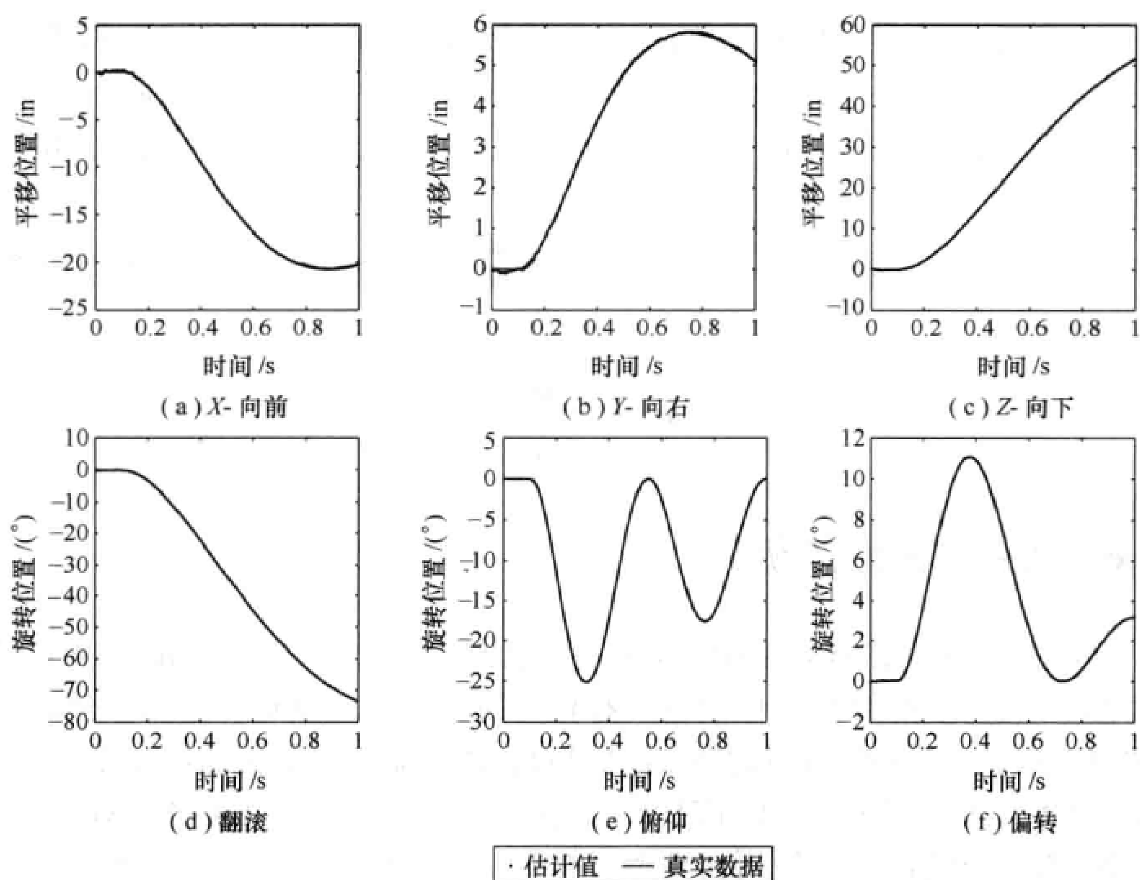


图 20－2 使用传感器融合估计器得的平移和旋转位置

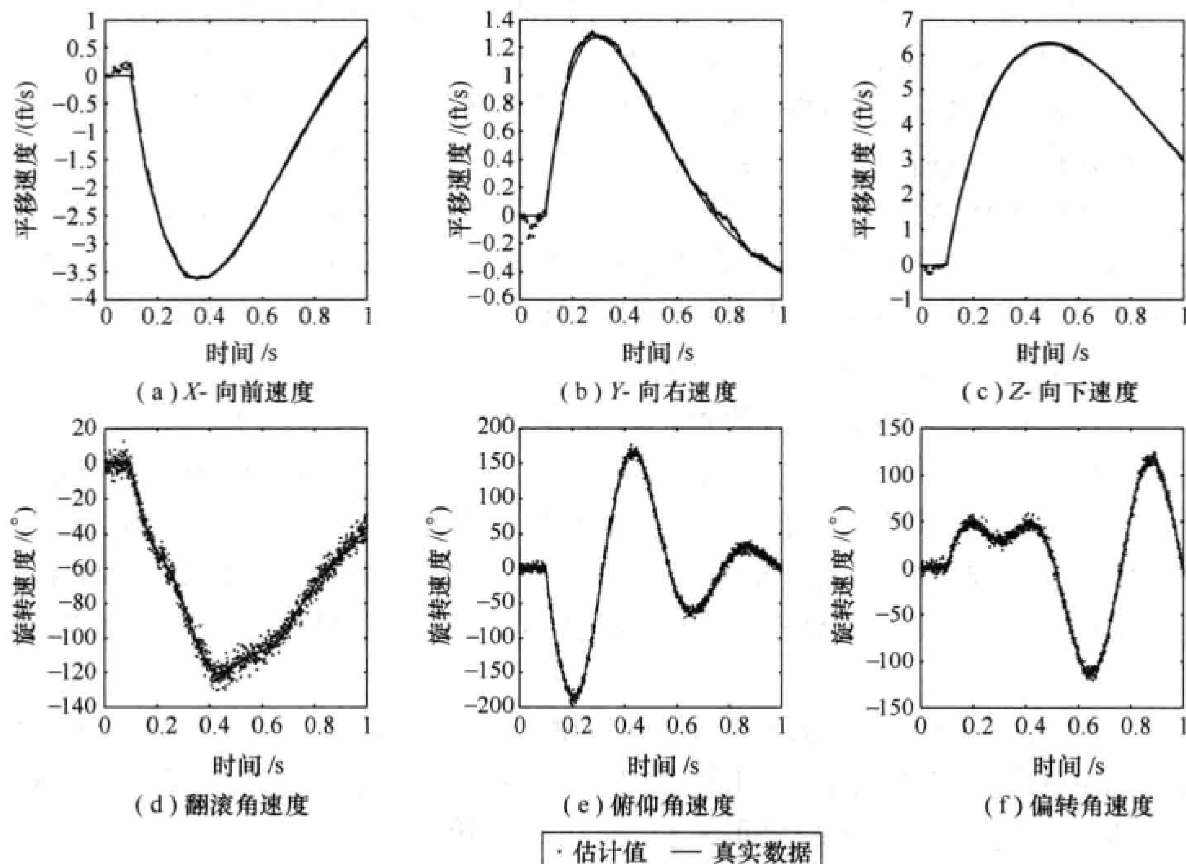


图 20－3 使用传感器融合估计器得的平移和旋转速度

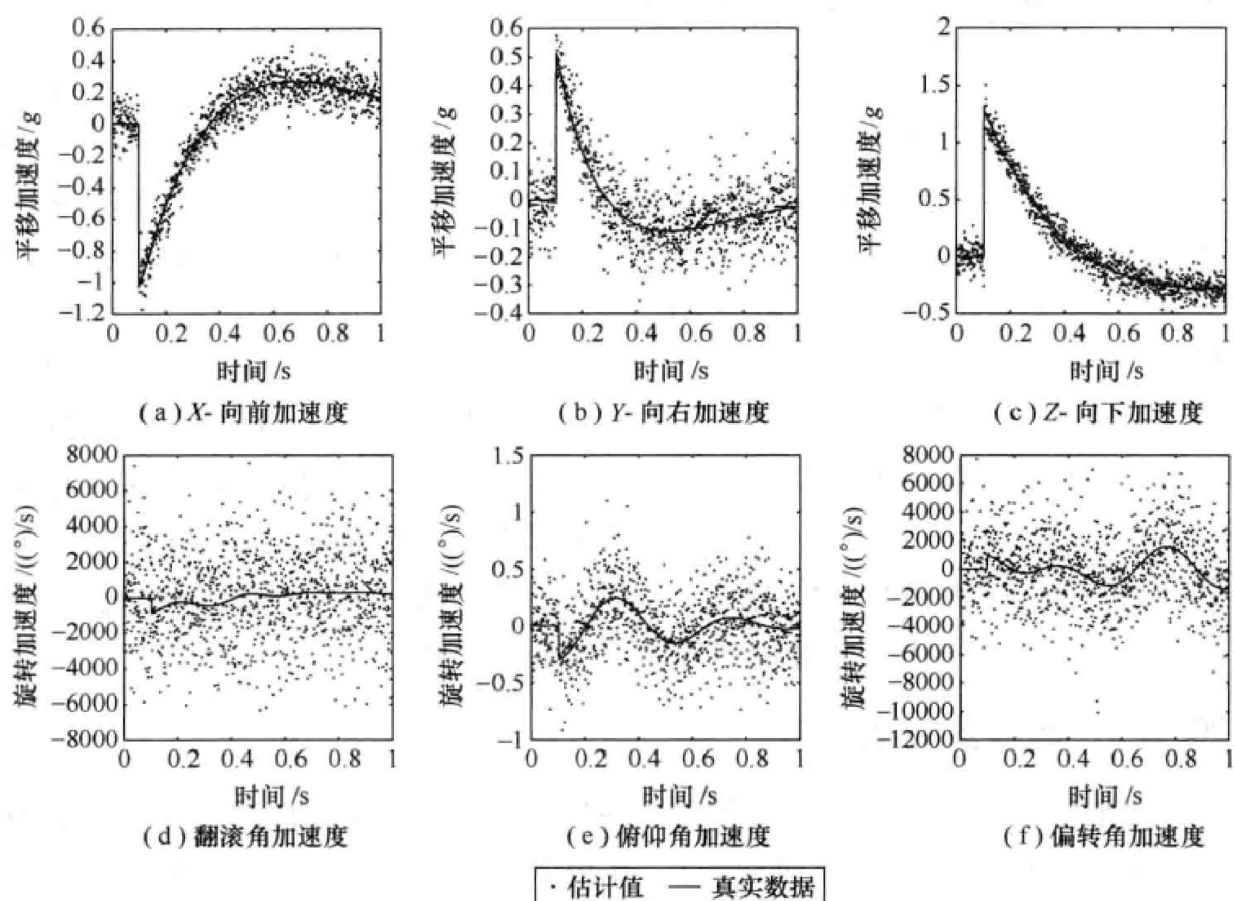


图 20－4 使用传感器融合估计器得的平移和旋转加速度

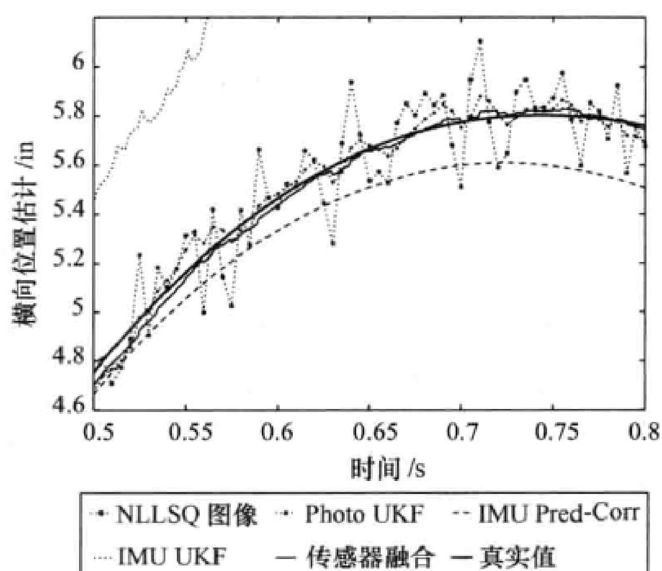


图 20－5 由多个估计过滤器使用相同的输入测量值得的横向位置估计

这些初步结论将在第 20. 6. 2 节进行检验，那里将给出各种估计器的均方根误差比较。

20.6.1 滤波器性能对比方法学

第 19.6.1 节提出一种性能评价方法,利用均方根误差衡量性能,这里再次使用该方法。同样,采用第 19.5 节提出的模拟无噪航迹 $\{\boldsymbol{x}_n, n = 0,1,2,\cdots,N\}$ 进行分析,而 $N = 1000$ 是第 20.4.1 节提出的。首先将航迹变换为一组无噪的模拟真实观测值 $\{\boldsymbol{h}_n(\boldsymbol{x}_n), n = 0,1,2,\cdots,N\}$,这里 $\boldsymbol{h}_n(\boldsymbol{x}_n)$ 可包含摄影测量法或 IMU 测量值,也可二者同时包含,用标记位 F_n 来进行不同情况的区分。$N_{\mathrm{RMS}} = 100$ 的观测噪声集合 $\{z_{n,r}^{0}, n = 0,1,2,\cdots,N; r = 1,2,\cdots,N_{\mathrm{RMS}}\}$ 是通过在模拟真实观测值中加入独立零均值高斯噪声产生的 N_{RMS} 含噪观测值集合。每个跟踪滤波器用一个含噪观测值集合用来训练,以生成 N_{RMS} 状态矢量估计集合 $\{\hat{\boldsymbol{x}}_{n|n,r}, n = 0,1,2,\cdots,N; r = 1,2,\cdots,N_{\mathrm{RMS}}\}$。计算每个跟踪滤波器的均方根误差相对时间的函数,得 3D 平移和角度位置误差 $\boldsymbol{e}_{\boldsymbol{p},n}^{(\mathrm{filter})}$ 和 $\boldsymbol{e}_{\boldsymbol{a},n}^{(\mathrm{filter})}$、3D 平移和角速度误差 $\boldsymbol{e}_{\dot{\boldsymbol{p}},n}^{(\mathrm{filter})}$ 和 $\boldsymbol{e}_{\dot{\boldsymbol{a}},n}^{(\mathrm{filter})}$ 以及 3D 平移和角加速度误差 $\boldsymbol{e}_{\ddot{\boldsymbol{p}},n}^{(\mathrm{filter})}$ 和 $\boldsymbol{e}_{\ddot{\boldsymbol{a}},n}^{(\mathrm{filter})}$。

20.6.2 滤波器对比结果

对于每个类型的估计器,计算均方根误差关于时间的函数,然后得出每种情况的平均和最大误差(一段时间内)。在表 20-1、表 20-2 和表 20-3 中给出了这些平均和最大值的概括总结,分别对应位置、速度和加速度均方根误差。

表 20-1 模拟数据 RMS 位置误差

滤波器类型	数据类型	位置/in		方向/(°)	
		均值	极大值	均值	极大值
NLLSQ	Photo	0.68	0.81	0.22	0.26
UKF(2)	Photo	0.32	0.52	0.15	0.22
Pred/Corr	IMU	1.21	3.09	0.55	0.85
UKF(2)	IMU	1.50	5.15	1.69	2.10
UKF(2)	Both	0.16	0.42	0.06	0.22

表 20-2 模拟数据 RMS 速度误差

滤波器类型	数据类型	平移速度/(ft/s)		角速度/((°)/s)	
		均值	极大值	均值	极大值
NLLSQ	Photo	N/A	N/A	N/A	N/A
UKF(2)	Photo	0.71	1.63	12.34	22.18
Pred/Corr	IMU	0.30	0.48	8.66	9.75
UKF(2)	IMU	0.41	1.05	7.47	8.55
UKF(2)	Both	0.10	0.24	7.43	8.53

表 20－3　模拟数据 RMS 加速度误差

滤波器类型	数据类型	平移加速度/g		角加速度/((°)/s^2)	
		均值	极大值	均值	极大值
NLLSQ	Photo	N/A	N/A	N/A	N/A
UKF(2)	Photo	0.37	1.70	653.3	3005.9
Pred/Corr	IMU	1.7	0.2	N/A	N/A
UKF(2)	IMU	0.16	0.24	4532.1	5106.0
UKF(2)	Both	0.15	0.89	4217.1	5119.2

图 20－6、图 20－7 和图 20－8 分别给出五种测试估计器中的每一种的时程：位置和旋转 RMS 误差、位置和旋转速度 RMS 误差以及位置和旋转加速度 RMS 误差。由表 20－1～表 20－3 和图 20－6～图 20－8 可见，除旋转加速度情况以外，使用摄影测量和 IMU 传感器数据融合的 UKF(2) 估计器的性能最好。

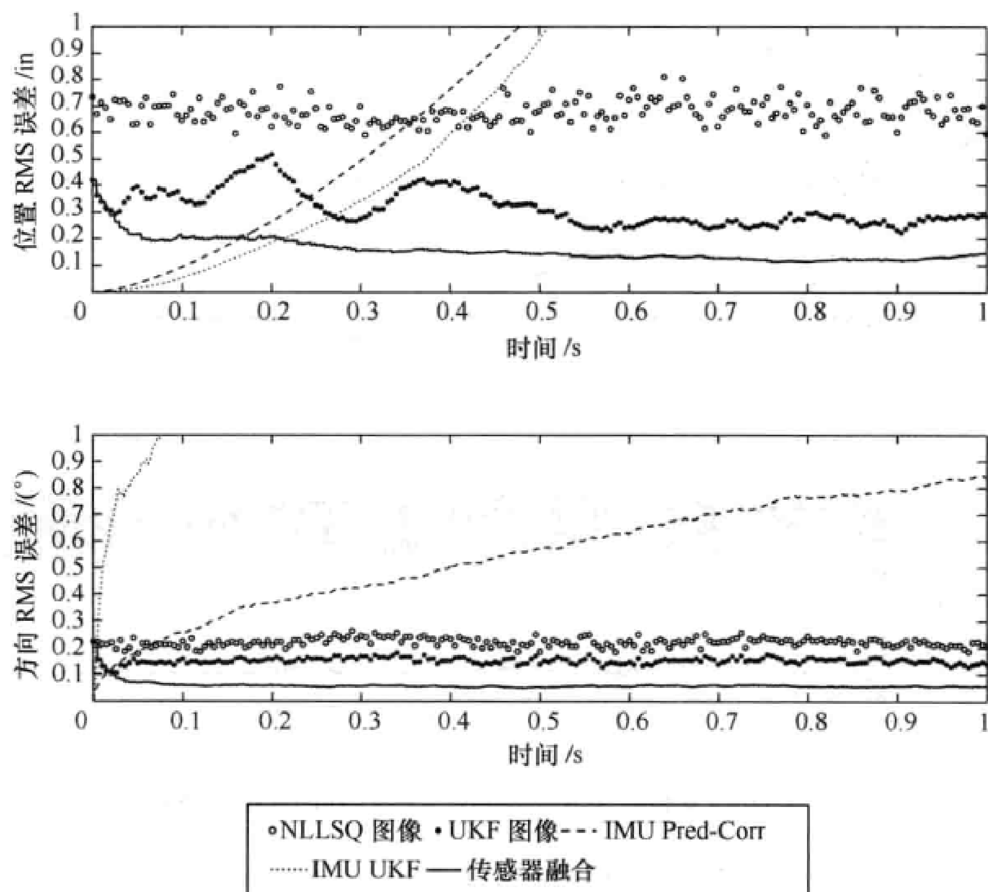


图 20－6　所有跟踪滤波器使用模拟数据得的位置和方向的 RMS 误差

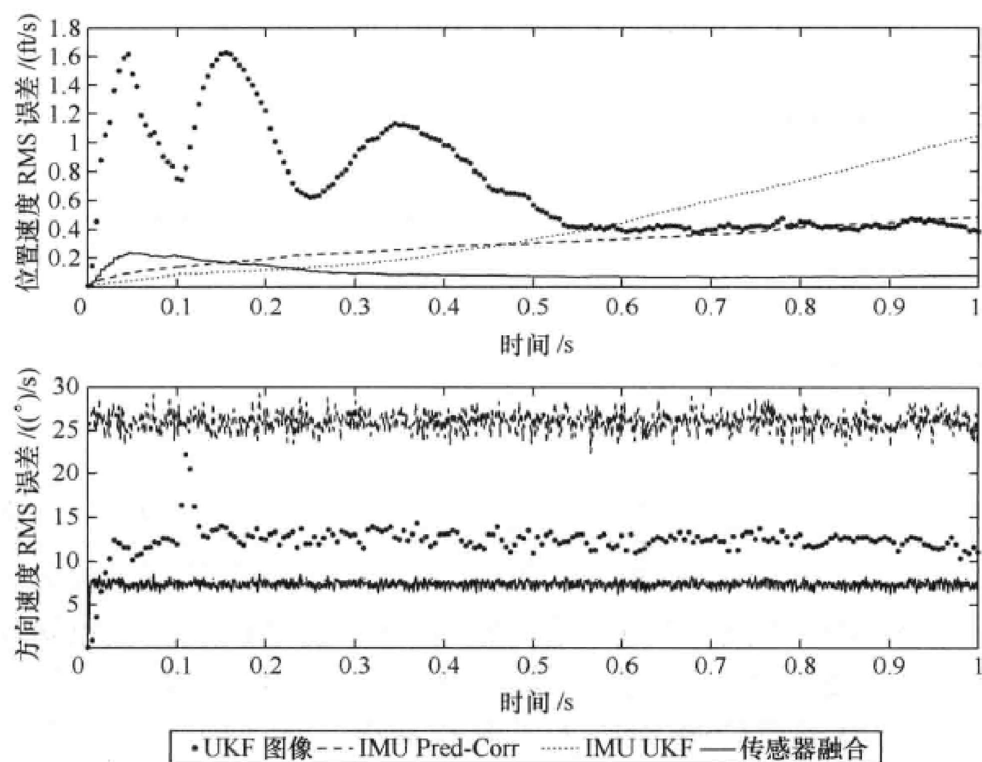


图 20－7　所有跟踪滤波器使用模拟数据得的位置和方向速度的 RMS 误差

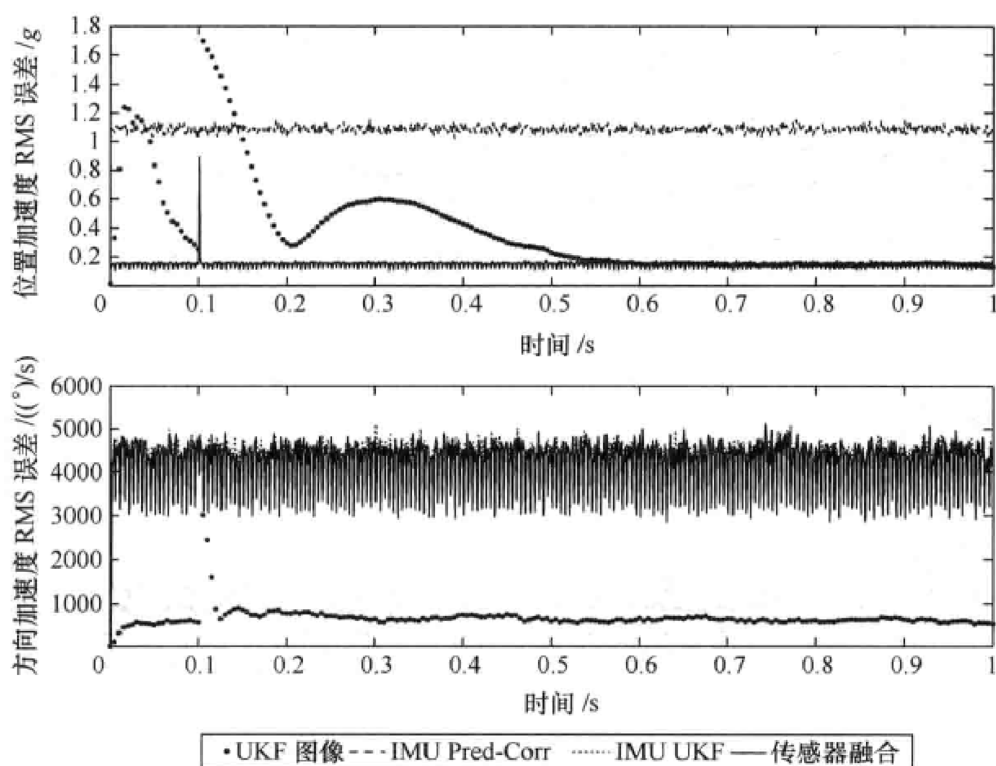


图 20－8　所有跟踪滤波器使用模拟数据得的位置和方向加速度的 RMS 误差

20.7 结 论

本章已经证明了传感器融合估计器得的估计结果比单独基于摄影测量或基于 IMU 的估计更为准确。在 UKF 求解器中,两种测量值互为补充。

IMU 测量值减轻摄影测量数据中的噪声,从而使位置和角度更为准确;而摄影测量值补偿了 IMU 测量值的偏差和漂移,使平移速度更为准确。

20.8 展 望

当前的测量模型假定 IMU 安装在刚体坐标系重心(CG)。为增加传感器融合估计的实用性,应该在测量模型中增加杆臂校正,以弥补加速度计中由刚体重心旋转所导致的向心力和切向线性加速度。本章还试图证明,使用传感器融合分析摄影测量和 IMU 数据,会提高容错,一种数据源中漏掉的部分将会被另外数据源的测量值所弥补。

除了可提高精度和容错性,传感器融合还可以提高摄影测量周转时间。在摄影测量数据简化过程中,二维跟踪数据的收集是最耗时的一步。本章试图通过使用传感器融合重新组织这一过程,以减少为得足够精确的 6 自由度解所需的图像测量值的数量。UKF 的不确定性模型提供了特定的事件,对 6 自由度解的准确性进行衡量。可以将传感器融合结果的不确定性与仅采用摄影测量的结果进行比较,以确定能获得足够的精度情况下,可收集的最少的摄影测量数据。

尽管当前 IMU 中速率陀螺仪似乎很准确,但加速度计的校准偏差却很大,会导致 6 自由度解的严重错误。为进一步提高精度,使用传感器融合来估计校准偏差。同时,假定在物体与飞机距离很近的短暂时间内,飞机所承受的力为常数。结合飞机 INS 系统数据,来确定校正飞机的加速度和旋转速率的变化是否会导致结果明显的差异。

本章还试图解决 IMU 数据延迟的问题。用于摄影测量的图像是带有收集和记录到飞机上的时间标记的,IMU 数据从目标传输到地面站,在加上时间标记之前,都收集在一个数据缓冲之中。这些多重数据通道导致不确定长度的时间同步误差。当允许不同延迟时,本章试图通过检查传感器融合解的残留误差来估计延迟。

参考文献

[1] lu SL, Wohn K. Estimation of General Rigid Body Motion from a long Sequence of Images. Technical Report, Department of Computer Science, University of Pennsylvania, 1990.

[2] Lee JW, Kim MS, Kweon IS. A Kalman filter based visual tracking algorithm for an object moving in 3D. *IROS' 95 Proceedings International Conference on Intelligent Robots and Systems*, 1995, 1.

[3] Gennery DB. Visual tracking of known three – dimensional objects. *Int. J. Comput. Vision*, 1992, 7(3): 243 – 270.

[4] Huang TS, Netravali AN. Motion and structure from feature correspondences: a review. *Proc. IEEE*, 1994, 82(2): 252 – 268.

[5] Ude A. Filtering in a unit quaternion space for model – based object tracking. Robot. Au – tonom. Syst, 1999, 20: 163 – 172.

[6] Halvorsen K, Soderstrom T, Stokes V, and Lanshammer H. Using an extended Kalman filter for rigid body pose estimation. J. Biomech. Eng, 2005, 127: 475 – 483.

[7] Algrain MC, Saniie J. Estimation of 3D Angular motion using gyroscopes and linear accelerometers. *IEEE Trans. Aero. Electron. Syst*, 1991, 27(6): 910 – 920.

[8] Marins JL, Yun X, Bachmann ER, et al. An extended Kalman filter for quaternion – based orientation estimation using MARG Sensors. In *Proceedings of the* 2001 *International Conference on Intelligent Robots and Systems*, 2001: 2003 – 2011.

[9] Yun X, Lizarraga M, Bachmann ER, et al. An improved quaternion – based filter for real – time tracking of rigid body orientation. In *Proceedings of the* 2003 *International Conference on Intelligent Robots and Systems*, 2003: 1074 – 1079.

[10] Sabatini AM. Quaternion – based extended Kalman filter for determining orientation by inertial and magnetic sensing. *IEEE Trans. Biomech. Eng*, 2006, 53(7): 1346 – 1356.

[11] Yun X, Bachmann ER. Design, implementation, and experimental results of a quaternion – based Kalman filter for human body motion Tracking. *IEEE Trans. Robot*, 2006, 22(6): 1216 – 1227.

[12] Kelly J, Sukhatme GS. Fast relative pose calibration for visual and inertial sensors. *Proceedings of the* 11*th International Symposium on Experimental Robotics* (*ROBIO*' 08), 2008.

[13] Bleser G and Strieker D. Advanced tracking through efficient image processing and visual – inertial sensor fusion. *Comput. Graphics*, 2009, 33: 59 – 72.

[14] Jeon S, Tomizuka M, Katou T. Kinematic Kalman filter (KKF) for robot end – effector sensing. *J. Dyn. Syst. Meas. Control*, 2009, 021010: 8.

[15] *Moler C.* Numerical Computing with Matlab. 2*nd ed. SIAM*, 2004.

[16] Van Loan CF. *Introduction to Scientific Computing.* 2nd ed. Prentice Hall, 2000.